

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 24, 2002, 05:31:50 ; Search time 2313.83 Seconds  
(without alignments)  
16279.401 Million cell updates/sec

Title: US-09-758-269-5  
Perfect score: 1800  
Sequence: 1 atggctttcttcacgcaac.....tggcgaagcaggctcgtgtga 1800

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues  
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_hg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_ats.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vl.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_ats.\*  
28: em\_un.\*  
29: em\_vl.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htgo\_inv.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
------------	-------------	-------	--------	----	-------------

ALIGNMENTS

RESULT 1

AX148306

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

thale cress.

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsia.

1 (bases 1 to 1800)

Iuchi, S., Kobayashi, M. and Shinozaki, K.

Transgenic plants carrying neoxanthin cleavage enzyme gene

Patent: EP 1116794-A 5 18-JUL-2001;

Riken (JP)

Location/Qualifiers

1..1800

/organism="Arabidopsis thaliana"

/db\_xref="taxon:3702"

1..1800

/note="unnamed protein product"

AX148306 Sequence

AB026549 Arabidops

AB028617 Arabidops

AY056255 Arabidops

AC013430 Genomic s

AX148316 Sequence

Z97215 Lycopersico

AXJ76244 Solanum t

AX148312 Sequence

AB030293 Vigna ung

AF190462 Phaseolus

AC074176 Arabidops

AX148302 Sequence

AL021710 Arabidops

AL161548 Arabidops

AF224672 Persea am

AF224671 Persea am

AX148314 Sequence

U95953 Zea mays vi

AX148310 Sequence

AB028621 Arabidops

AL021713 Arabidops

AF399655 Lactuca s

Z93765 M. domestica

AX148304 Sequence

AY056789 Arabidops

AL021687 Arabidops

AL161550 Arabidops

AX085456 Sequence

AY029525 Phaseolus

AP004129 Oryza sat

AX148308 Sequence

AX148318 Sequence

AJ005813 Arabidops

AF224670 Persea am

AJ416714 Crocus sa

AL663011 Oryza sat

AF324838 Streptomy

AJ416712 Crocus sa

AF322256 Streptomy

AC074054 Oryza sat

AC074355 Oryza sat

AE006980 Mycobacte

Z95210 Mycobacteri

AP004729 Oryza sat

AX148306 1800 bp

Sequence 5 from Patent, EP1116794.

AX148306

AX148306.1 GI:14347193

thale cress.

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsia.

1 (bases 1 to 1800)

Iuchi, S., Kobayashi, M. and Shinozaki, K.

Transgenic plants carrying neoxanthin cleavage enzyme gene

Patent: EP 1116794-A 5 18-JUL-2001;

Riken (JP)

Location/Qualifiers

1..1800

/organism="Arabidopsis thaliana"

/db\_xref="taxon:3702"

1..1800

/note="unnamed protein product"

```
/codon_start=1
/protein_id="CAC41194.1"
/db_xref="GI:14347194"
/translacion="MASFTATAVSGRWLGNHTQPLSSQSSDLSYCSSLPMASRV
TRKUNVSAHLTPPALHPFKOSSNPAIVVKPKAKESNTKOMNLFORAAALDAABG
FLVSEKHLPLKPTADPSVOTAGFAVNEQPVRRNLPVVGKLPDSIKGVYVRNGANP
LHEPVTGHHFDFDGMVHAKFHEGSAAYACRFQTQNRFOERQLGRPFPPKAI GEIH
GHTGIAELMLFYARAAAGIVDPAHGTCVANAGLVYFNGRLLAMSEDDLPHYVOITPNG
DLKTVGRFDDQLESTWIAHPKVDPSGELFALSVDVSKPLYKIFRFPSPDGTKSPD
LVEIDQPTWMHDAITENFVVPDQVFKLPEMIRGSPVVDKNKXVARFGLLDKY
AEDSSNKWIDAPDFCFHLMAWEEPEFDEVVIGSCMTFPPDSIFNESDENLKSVLS
EIRLNLTGSGSTRPIISNEDQVNLBGMVNRMLGRKTKFAYLALAEWPVKVSGFA
KVDLTTCGEVKHLYGDMRKYGEPLFLPGEGEDEGEVILCFVHDEKTKWSELQIVNAV
SLEVEATVKLPSRVPYGFHGTFFGADDLAKQVV"

BASE COUNT      458 a  464 c  439 g  439 t
ORIGIN

Query Match      100.0%; Score 1800; DB 6; Length 1800;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1800; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atggctttttcacggcaacggctggctttctggagatggcttggcggaatcatact 60
Db 1 ATGGCTTCTTTCACGGCAACGGCTCGGCTTCTGGAGATGGCTTGGTGGCAATCATACT 60

Qy 61 cagcggccattatgctcttctcaaaagctccgacttgatgtattgttagtctcttaccatg 120
Db 61 CAGCGCCATTATGCTCTTCTCAAAAGCTCCGACTTGAATTATGTAGTCTCTTACCTATG 120

Qy 121 gccagtcgtgcacagcaagctaaatttccatctcgcttccacactcctccagctctt 180
Db 121 GCCAGTCGTGTGCACACGTAAGCTCAATGTTTCATCTCGCTTCACACTCTCCAGCTCTT 180

Qy 181 catttccctaaagcaatcatcaaaactctccgccattgtgttaagcccaagccaaagaa 240
Db 181 CATTTCCCTTAAGCAATCATCAAACTCTCCGCCATTGTTGTTAAGCCCCAAAGAA 240

Qy 241 tccaaactaaacagatgaatttgttcagagagcggcggcggaagcttggagcggcg 300
Db 241 TCCAACTAAACAGATGAATTTGTTCAGAGAGCGCGCGGCGGCGGCTTGAGCGCGCG 300

Qy 301 gaggttccctgtcagccacgagagctacaccgccttcttcaaacggctgacacctagt 360
Db 301 GAGGTTTCTTGTTCAGCCACGAGAAAGCTACACCCGCTTCTTAAACCGGCTGATCCTAGT 360

Qy 361 gttcagatcgccggaaattttgtctccggtgaatgaacagccgctccggcgtaattctccg 420
Db 361 GTTCAGATCGCGGAAATTTGTCTCCGGTGAATGAACAGCCGCTCCGCGGTAACTCTCCG 420

Qy 421 gtggtcggaaaaacttcccgattccatcaaaagagtgtatgtgcgaacggagctaaacca 480
Db 421 GTGGTCGGAATAACTTCCCGATTCCATCAAAAGAGTGTATGTGCGCAACGGAGCTAAACCCA 480

Qy 481 cttcacgagcgggtgacaggtcacactcttccagcagagagctatggttcacgcgctc 540
Db 481 CTTACAGAGCGGTTACAGGTACACACTTCTTCAGCGAGACGGTATGGTTACACGCCGTC 540

Qy 541 aaattcgaacacgggttcagctagctacgcttgcgggttactcagactaaacgggtttgt 600
Db 541 AAATTCGAACACGGTTACAGTACGCTTGCCTGGTTCCTTACTCAGACTTAAACGGTTTGT 600

Qy 601 caggaaactcaattgggtcgaccggttttccccaagccatcggtgagcttcacggccac 660
Db 601 CAGGAACGTCAATTGGGTGCGACCGGTTTTTCCCAAGGCCATCGGTGAGCTTACCGGCCAC 660

Qy 661 accgattgcccactcatgctattcacccagagctgagccggtatagtcgaccgcg 720
Db 661 ACCGATTGCCCCGACTCATGCTATTCTACGCCAGAGCTGACGCCGGTATAGTCACCCCG 720

Qy 721 gcacagcaacgggtgtagctaacccgggtttgtgctatttcaatggccggctattggct 780
Db 721 GCACAGCAACGGGTGAGCTAAACCGGTTTGGTCTATTTCATATGGCCGGTTATTGGCT 780
```

```
Qy 781 atgtcggagagatgattttacattaccgaattccagactcactccccaatggagatttataaac 840
Db 781 ATGTCCGAGGATGATTTTACCTTTCAAGTTTCAGATCACTCCCAATGAGATTTAAAAACC 840

Qy 841 gttggtcggttcgatttggatggacaattagaatcccaaatgattgccaccgccgaagtc 900
Db 841 GTTGTGCGTTCGATTTTGTATGGACAATTAGATCCACAATGATTGCCACCCGAAAGTC 900

Qy 901 gaccggaaatccggatgaactcttgcgttttaagctacagctcgctttcaaaagccttacct 960
Db 901 GACCCGGAATCCGGTGAACCTCTTCGCTTTTAAGCTACACGCTCGCTTTCAAAAGCCTTACCTA 960

Qy 961 aaatacttcgattctcacggagcgaactaaatacacggagcgtcgagattcaacttgat 1020
Db 961 AAATACTTCCGATTCTCACCGGACGGAACTAAATCACCGGACGTCGAGATTCACGCTTGAT 1020

Qy 1021 cagccaaacgatgatgcacgatttcgcgatttacagagaacttcgtcgtacctgaccag 1080
Db 1021 CAGCAACGATGATGCACGATTTCCGATTTACAGAGAACTTCGTCGTCGTACCTGACCCAG 1080

Qy 1081 caagtcgttttcaagctgccggagatgatccgggtgggttcccggtgggttttacgacaag 1140
Db 1081 CAAGTCGTTTTCAAGCTGCCGAGATGATCCGCGTGGGTCTCCGCTGGTTTTACGACAAG 1140

Qy 1141 acaaggtcgcaagatttcgggatttttagacaaataacgccaagatttcataaataaag 1200
Db 1141 AACAGGTCCGAAGATTCGGGATTTTAGACAAATACGCCGAAGATTCATCGAAATTAAAG 1200

Qy 1201 tggattgatgctccagattgcttctgcttccatctctggaaacgcttgggaagagccagaa 1260
Db 1201 TGGATTGATGCTCCAGATTGCTTCTGCTTCCATCTCTTGAAACGCTTGGGAAGAGCCAGAA 1260

Qy 1261 acagatgaagtgcgtgtagagggtcctgtagctccacagactcaattttcaacgag 1320
Db 1261 ACAGATGAAGTCGTCGTGATAGGTCCTGTATGATCTCCACAGACTCAATTTTCAACGAG 1320

Qy 1321 tctgacagagaatctcaagagtgctctgtctgaaatcgcctgaatctcaaaacgggtgaa 1380
Db 1321 TCTCAGCAGAAATCTCAAGAGTGTCTGTCTGAAATCGCCCTGAATCTCAAAACCCGGTAA 1380

Qy 1381 tcaactcgccgtccgactcatctccaaacgaagatcaacaagtcaacctcgaagcaggagat 1440
Db 1381 TCAACTCGCCGTCCGATCATCTCCACGAGATCAACAAGTCAACCTCGAAGCAGGGATG 1440

Qy 1441 gtcaacgagaacatgctcggcgttaaaccaaatctgcttacttctttagcgcgagccg 1500
Db 1441 GTCAACGAGAAACATGCTCGGCCGTAAACCAAAATTCGCTTACTTGGCTTTAGCCGAGCCG 1500

Qy 1501 tggcctaaagtctcaggatttcgctaaagttagatctcaactactggagaagtaagaacat 1560
Db 1501 TGGCCTAAAGTCTCAGGATTTCGCTAAAGTTGATCTCACTACTGGAGAAAGTTAAAGAAAT 1560

Qy 1561 cttcacggcgataaccgttacggagggagagcctctgttctccccggagagagggagag 1620
Db 1561 CTTTACGGCGATAACCGTTACGGAGGAGAGCCCTCTGTCTTCTCCCGGAGAGGAGGAGAG 1620

Qy 1621 gaagacgaaggatacatcctctgttttcggttcacgacgagaagacatggaatcgagttta 1680
Db 1621 GAACACGAAGGATACATCTCTGTTTCGTTACGACGAGAAGACATGGAATTCGAGTTA 1680

Qy 1681 cagatagtttaacggcggttagcttagaggttgaagcaacgggttaaaacttcgtaagggtt 1740
Db 1681 CAGATAGTTTAAACGGCGTTAGCTTAGAGGTTGAAGCAACGGTTAAACTTCGCTCAAGGGTT 1740

Qy 1741 ccgtacggatttcacggtacattcatcgagccgagctgatttggcgaagcaggtcggtgga 1800
Db 1741 CCGTACGGATTTACGGGTACATTTCATCGAGCCGATGATTTCGCGAAGCAGGTCTGTGTGA 1800

RESULT 2
AB026549 3869 bp DNA linear PLN 24-OCT-2001
LOCUS
```

```

DEFINITION  Arabidopsis thaliana gene for neoxanthin cleavage enzyme, complete
              cds.
ACCESSION  AB026549
VERSION    AB026549.1  GI:16416373
KEYWORDS   neoxanthin cleavage enzyme.
SOURCE     Arabidopsis thaliana (strain:Col.) DNA.
ORGANISM   Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eustoids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE  1 (bases 1 to 3869)
AUTHORS   Iuchi,S., Kobayashi,M. and Shinozaki,K.
TITLE      Characterization of neoxanthin cleavage enzyme from Arabidopsis
            thaliana
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 3869)
AUTHORS    Iuchi,S. and Shinozaki,K.
TITLE      Direct Submission
JOURNAL    Submitted (22-APR-1999) Satoshi Iuchi, RIKEN, Plant Mol. Bio.;
            3-1-1 Koyadai, Tsukuba 305-0074, Japan
            (E-mail:luchi@rtc.riken.go.jp, Tel:81-298-36-4359)
FEATURES   Location/Qualifiers
            1..3869
             /organism="Arabidopsis thaliana"
             /strain="Col."
             /db_xref="taxon:3702"
            1831..3830
             /codon_start=1
             /product="neoxanthin cleavage enzyme"
             /protein_id="BAB70609.1"
             /db_xref="GI:16416374"
             /translations="MASFTATAASQRWGLGNHTQPLSSQSSDLSSYSSSLPMASRV
            TRLVNSALLHTPALHPKQSSPAIVVKPKASNTKOMLRFORAAAALDAAGP
            FLYSHKELPLTPADSPVOIAGNPVNEQPVRLPVVGLPDSTKGVVVRNGAP
            LHPVTHGHPFDGDMVHAKVPHGSASYACRTQTNRFVQRQLGRPVPKAIGELH
            GHTGARLMULFYARAAGAVDPAHGTGVAGNGLVYNGRLUWSEDDLPQVOITPNG
            DLKTGRFDGGLGSELTMIADKVPDSEGLFALSVDVWSKPLYKFKRPSDGTGKPD
            VETLQDPTMHDFAITENFVNVQDQVFKLPMIRGGSPVYDKNVARFGILDKY
            AEDSSNKLTDASDCIFCFHLNNAWEPETDEVVVI GSCMTPPDSIFNESDENLKSLS
            EIRLNLTGSTRPILISNEDDQVNLNAGMVNENMLGRKTKFAYLALAPMPKVSQFA
            KVDLTWTGCEKHLIYGNRYGGEFLPLPBGEGEDEGYILCFVHDEKTKWSELQIVNVA
            SLEVTAEKVLPSRPVPGYGFIFGADDLAKQVV"
BASE COUNT 1155 a   845 c   773 g   1096 t
ORIGIN
Query Match      100.0%; Score 1800; DB 8; Length 3869;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1800; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  1  atggctcttctcagcgcaacggtcgcgggttctctggagatggcttgatggcaatcatact 60
      |||||
Db   1831  ATGGCTTCTTTCACGGCAACGGCTCGCGTTTCTGGGAGATGGCTTGTGGCAATCACTACT 1890
      |||||

Qy  61  cagcgccattatcgctcttctcaaagctccgacttgattttagctctcttaccctatg 120
      |||||
Db   1891  CAGCGCCATTATCGTCTTCTCAAAGCTCCGACTTGAGTTATTGTAGCTCTTACCTATG 1950
      |||||

Qy  121  gccagtctgtccacacgaagctcaattgtttcattcgcttccacctctccagctctt 180
      |||||
Db   1951  GCCAGTCTGTGCACACGTAAGCTCAATGTTTTCATCTCGCTTTCACACTCTCTCCAGCTCTT 2010
      |||||

Qy  181  catttccctaagcaatcatcaactctccgcgaattgtttaaagcccaagcaaaagaa 240
      |||||
Db   2011  CATTTCCCTTAAGCAATCATCAACTCTCCCGCCATTGTTGTTAAGCCCAAGCAAGAA 2070
      |||||

Qy  241  tccaactactaaacagatgaatttttccagagagcggcgcgagcgttggagcgcgcg 300
      |||||
Db   2071  TCCAACACTAAACAGATGAATTGTTTCCAGAGACGCGCGCGGAGCGTTGGACGCGCGG 2130
      |||||

Qy  301  gagggttctcttgcagccacgagaagctacacccgcttctctaaacgggtgatcctagt 360
      |||||
Db   2131  GAGGGTTCTCTTGTGACGCACGAGAAGCTACACCGCTTCTCTAAAACGGGTGATCTAGT 2190
      |||||

```

Qy	1441	gtcaacagaaacatgctcgccgtaaaccaaaatcgcttacttgctttagccgagccg	1500
Db	3271	gtcaacagaaacatgctcgccgtaaaaacaaatcgcttacttgctttagccgagccg	3330
Qy	1501	tggcctaagtcctcaggattcgctaaagttagcttcactactgagaaagttaagaaacat	1560
Db	3331	tggcctaaagtctcaggattcgctaaagttagcttcactactgagaaagttaagaaacat	3390
Qy	1561	ctttaacgcgataaacctgtaacgaggaagacctctgtttctcccggaagaaggagagag	1620
Db	3391	ctttacgcgcgataaacctgtaacgaggaagacctctgtttctcccggaagaaggagagag	3450
Qy	1621	gaagcgaagatgatacatcctctgtttctcagcagcagaagacatgaaatcggagatta	1680
Db	3451	gaagcgaagatgatacatcctctgtttctcagcagcagaagacatgaaatcggagatta	3510
Qy	1681	cagatagttaacgcgcgttagcttagagagttgaagcaaacggttaaaacttcgcgtcaagggtt	1740
Db	3511	cagatagttaacgcgcgttagcttagagagttgaagcaaacggttaaaacttcgcgtcaagggtt	3570
Qy	1741	ccgtacgagattcaacggttacattcatcgcagcagcagatgattggcgaacagagtcgtgtga	1800
Db	3571	ccgtacgagattcaacggttacattcatcgcagcagcagatgattggcgaacagagtcgtgtga	3630
RESULT 3			
AB028617/c			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Arabidopsis thaliana			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Residae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (sites)			
Sato, S., Nakamura, Y., Kaneko, T., Katoh, T., Asamizu, E. and Tabata, S. Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence features of the regions of 4,504,864 bp covered by sixty P1 and TAC clones			
DNA Res. 7 (2), 131-135 (2000)			
20277480			
2 (bases 1 to 52232)			
Sato, S., Nakamura, Y., Kaneko, T., Kato, T., Asamizu, E. and Tabata, S. Submitted (09-JUN-1999) Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan [E-mail: ynakamu@kazusa.or.jp, Tel:81-438-52-3935, Fax:81-438-52-3934]			
Address for correspondence: kaos@kazusa.or.jp			
For the latest information on annotation of this clone, please see http://www.kazusa.or.jp/kaos/cgi-bin/agd_graph.cgi?c=MOA2			
Genes with similarity to proteins in the databases are described in 'product or 'note' qualifiers. Genes that have no significant protein similarity are described as 'unknown protein'.			
The software programs used to predict genes include: Grail (Informatics Group, Oak Ridge National Laboratory, http://compbio.ornl.gov/Grail-1.3/), GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html), NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and SplicePredictor (Volker Brendel, Stanford University, http://gremilni.zool.iastate.edu/cgi-bin/sp.cgi).			
Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, http://genome.wustl.edu/eddy/tRNAscan-SE/).			
This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions. The 5' clone is MLN21 and the 3' clone is MIE1.			
Location/Qualifiers			
FEATURES			

source	
CDS	
CDS	
CDS	
CDS	

1. .52232	/organism="Arabidopsis thaliana"
	/strain="Columbia"
	/db_xref="taxon:3702"
	/chromosome="3"
	/clone="MOA2"
	/clone_lib="Mitsui P1"
join(1996. .2132,2225. .2349,2455. .2501,2648. .2725,	
2807. .2931,3021. .3087,3216. .3278,3370. .3439,3529. .3629,	
3718. .3822,3906. .4091)	
/notes="gene_id:MOA2.2"	
/codon_start=1	
/evidence=not experimental	
/product="glycolate oxidase"	
/protein_id="BAB01334.1"	
/db_xref="GI:11994212"	
/translation="MEITNVTEYDAIAKOKLPMVVDYVASGAEDOWTIOENRNAPAR	
IFRPRIIDVSKIDMTTIVLGFKISMPIMVAPTAKMWHPDGEYATRAAAGAT	
MTLSSWATSVSEVATPGIRFOLYVYKRNVDVQLVRAERAGFKAIALVTDTPR	
LGRRESDIKNRFTLPNLTAKNFEGLDLGKMDANDGLASYVAGOIDRTLSKDVQM	
LQITIKPLIVKGLTGEDARIAIQAGAAGIYVNHGARQLDYVPATISALEVVKAT	
QGRIPVLDGGVRGTDVFKALGASGIFGRPVVFSLAAGEAGVRKVLQMLRDEF	
ELTMALSGCRSLKEISRNIHTTEWDTTPRSARL"	
complement(join(4591. .4659,4763. .4924))	
/notes="gene_id:MOA2.3	
unknown protein"	
/codon_start=1	
/evidence=not experimental	
/protein_id="BAB01335.1"	
/db_xref="GI:11994213"	
/translation="MIGRRAGTNRVGMRRDSDLSLTRFVDSVFYFRFLAEPEILVFLFM	
LITVIFKDLTSRPNRILIVKEPKGSDIWF"	
complement(11870. .13669)	
/notes="gene_id:MOA2.4"	
/codon_start=1	
/evidence=not experimental	
/product="9-cis-epoxycarotenoid dioxygenase"	
/protein_id="BAB01336.1"	
/db_xref="GI:11994214"	
/translation="MASPTATAAVSGRWLGNTQPLSSQSSDLSYCSSLPMSARV	
TRKLNVSALHTPPALHPFKQSNPAIVVKPKAKSNTKQMLFQRAAAALDAABG	
FLVSHKLPKLPKTAADPSVQIAGNFAPVNEQPVRRNLPPVVKGLPDSIKGYVVRGANGP	
LHEPVTGHHFFDGDGMVHAVKFEHGSASACRFTQTNRFQERQLGRVPVPAIGELH	
GHTGIARMLFYAAAGIVDPAHGTGVANAGLVYFNGLLMSDDLPRYOVIITPNG	
DLTVGRFPDQGLSTMAHPKVDPESEGFALSVDVSKVLYKFRSPDGTGSPD	
VEILOQPTMMHDFALTENFVPPDQVVFVKLFEMIRGSGSPVYDKNKVARFGLDKY	
AEDSSNKWIDAPCFCHLWNAEETDDEVVIVIGSTPTPDSINENESDNKSVLS	
EIRLNLTGSETRRPIISNEDQQVLEAGMNRNMLGRKTKFAYLALAEPPKVGFA	
KVDLTITGEVKKHLYGDNRYGGEPLFLPBGGEDEGYILCFVHDEKTKWSELQIVNAV	
SLEVEATVKLPSRPVYGFHGTFTIGADDLAKQVY"	
join(29980. .30324,30421. .30528,30616. .30660,30841. .30913,	
31021. .31133,31239. .31298,31376. .31492,31588. .31649,	
31733. .31793)	
/notes="contains similarity to RNA-binding protein	
gene_id:MOA2.5"	
/codon_start=1	
/evidence=not experimental	
/protein_id="BAB01337.1"	
/db_xref="GI:11994215"	
/translation="MATVTMPDTPVVVDVDDKEVSTKSEKIIDEGIEKSSITDSKET	
ESLDDMKLVAMFKLNPFAKEFFPSYDPPKKNVAKANQFLPADDFETTKQSGEE	
FSLDAKDDNTKRRKNSYSGRRLTGRIKSAOREDRIIRTVYVSDIDQSVTEGLACL	
FNGCQVQDCRICGDSPHSLRFAFVFDADQGAHEALS LGGTMLGYPVAVLPSKTAI	
LPNVGTFPLRSEDEREMCTIYCTNIDKKVSQADVRNFFGESACGEVPIRULLGQDH	
STRIAFVEFALDALSALNCGMVGSQPIRVSPSKTPVRPIRTPPPSTN"	
complement(32182. .36456)	
/notes="gene_id:MOA2.6"	
/codon_start=1	
/evidence=not experimental	
/product="disease resistance complex protein"	
/protein_id="BAB01338.1"	
/db_xref="GI:11994216"	
/translation="MANSYLSSCANVMVERINTSQBELVELCKGSSSALLKRLKVALV	





QY	901	gaccggaatccgggtgaactctcggtttaagctacagcgtcggtttcaaaagccttaccta	960
Db	12769	GACCCGAATCCGGTGAACCTCTTCGTTTAAAGTACAGCTGCTTTCAAGAGCTTACCTA	12710
QY	961	aaatactccgattctcacggagcgaactaaataacccgagcgtcgagattcagcttgat	1020
Db	12709	AAATACTTCGATTCTCACCGGACGGAACCTAAATACCGGAGCTCGAGATTGAGCTTGAAT	12850
QY	1021	cagccaacgatgatgcacgatttcgagattacagagaacttcgctgctgactgaccag	1080
Db	12649	CAGCCAACGATGATGCAGGATTTCCGGATTACAGAGAACTTCGTCGTGCTGACTGACCAG	12590
QY	1081	caagtctgtttcaagctgcccggagatgacccggctgggtctccggtgggtttacacaga	1140
Db	12589	CAAGTCGTTTTCAAGCTGCCGGAGATGATCCGGCGTGGGTCTCCGGTGGTTTACGACAAG	12530
QY	1141	aaacaagtcgaagattcgggattttagacaaataacccgaagattatcgagaattaaag	1200
Db	12529	AACAAGTTCGAAGATTCGGGATTTTAGACAATAATACGCCGAAGATTTCATCGAACATTAAG	12470
QY	1201	tgattgatgtccagattgctctctctccatctctggaacgcttgggaagaccagaa	1260
Db	12469	TGGATTGATGCTCCAGATTGCTTCTGCTTCCATCTCTCGAAACGCTTGGGAAGACGACAA	12410
QY	1261	acagatgaagtcgctgtagaggttcctgtatgactccaccagactcaattttcaacgag	1320
Db	12409	ACAGATGAAGTCGTGATGATGAGGTCTCTGATGACTCCACAGACTCAATTTTCAACGAG	12350
QY	1321	tcgacgagaatctcaagatgtctctgtctgaaatccgctgaatctcaaaaccggtgaa	1380
Db	12349	TCGTACAGAGAATCTCAAGATGTCTCTGCTGAAATCCGCTGAATCTCAAAACCCGGTGAA	12290
QY	1381	tcaactcgccgtcgatcatctccaacgaagatcaacaagtcaactcgaaagcgggatg	1440
Db	12289	TCAACTCCCGTCGATCATCTCCACGAGAGATCAACAGTCAACTCGAAGCAGGGATG	12230
QY	1441	gtcaacagaacatgtcggccgtcaaaaccaaattcgcttacttggttttagccgagcgg	1500
Db	12229	GTCAACAGAAACATGCTCGCCCGTAAACCCAAATTCGCTTACTTGCTTTAGCCGAGCGG	12170
QY	1501	tggcctaaagcttcaggattcgtaagtgtatctcaactactggagaagtttaagaaacat	1560
Db	12169	TGGCCTAAAGTCTCAGGATTTCGTAAGTTGATCTCACTACTCGGAAGTTAAGAAACAT	12110
QY	1561	ctttacggcataaccgttacgagagagcctctgttctcccggagaaaggagagag	1620
Db	12109	CTTTACGGCCATACCCGTTACGAGAGAGCCTCTCTTTCTCCCGGAAAGAGAGAGAG	12050
QY	1621	gaagacgaagatacatctctgttcggtcacgacgagaagacatggaatcggagttga	1680
Db	12049	GAAGACGAAGATACATCTCTGTTTCGTTACGACGAGACATGGAATTCGAGATTA	11990
QY	1681	cgatagttaacgcgttagcttagaggttgaagcaacgggttaaaacttcgccaagggtt	1740
Db	11989	CAGATAGTTAAACCCGTTAGCTTAGAGTTTGAAGCAACCGGTTAAACTTCCTCGTCAAGGGTT	11930
QY	1741	ccgtacggatttcacggtcatcttcagcagcgatgatttggcgaagcagctcgtgtga	1800
Db	11929	CCGTACCGGATTTACGGTACATTCATCGAGCGCGATGATTGCGGAAGCAGGTCGTGTGA	11870
RESULT	4		
LOCUS	AY056255	2331 bp	mRNA linear
DEFINITION	Arabidopsis thaliana putative 9-cis-epoxycarotenoid dioxygenase (MOA2.4/AT3g14440) mRNA, complete cds.		
ACCESSION	AY056255		
VERSION	AY056255.1		
KEYWORDS	FLI CDNA		
SOURCE	thale cress.		
ORGANISM	Arabidopsis thaliana		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		

```
misc_difference 1234
/ gene="MOA2.4/AT3g14440"
/ note="compared to genomic sequence resulting in an amino
acid sequence difference"
/ replace="g"
1923 - 2331
. / gene="MOA2.4/AT3g14440"
misc_difference 2059
/ gene="MOA2.4/AT3g14440"
/ note="not present in genomic sequence"
misc_difference 2312
/ gene="MOA2.4/AT3g14440"
/ note="not present in genomic sequence"
misc_difference 2315
/ gene="MOA2.4/AT3g14440"
/ note="not present in genomic sequence"
BASE COUNT 615 a 560 c 533 g 623 t
ORIGIN

Query Match 99.9%; Score 1798.4; DB 8; Length 2331;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1799; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 atggctttttcacggcaacgcgctcggtttctggagatggcttggtggcaatcatact 60
Db 123 ATGGCTTCTTTTTCACGGCAACCGCTCGGTTTCTGGGAGATGGCTTGGTGGCAATCATACT 182

Qy 61 cagcgcattatcgtcttctcaagctccgacttgagttattgtagctccttaactatg 120
Db 183 CAGCCGCCATTATCGTCTTCTCAAGCTCCGACTTGAGTTATTGTAGCTTCTTACCTATG 242

Qy 121 gccagtcgtgcacacgtaagctcaatgtttcatctgcgcttcacactcctccagctctt 180
Db 243 GCCAGTCGTGTCACACGTAAGCTCAATGTTTCATCTCGCTTCACACTCTCTCCAGCTCTT 302

Qy 181 catcttcctaagaacatcacaactctccgccattgttgtaagcccaaaagaa 240
Db 303 CATTTCCCTTAAGCAATCATCAAACTCTCCGCCATTGTTGTTAAAGCCCAAGCCAAAGAA 362

Qy 241 tccaactaaacagatgaattgttcagagagcggcgcgagcgttgacgcggcg 300
Db 363 TCCAACACTAAACAGATGAATTGTTTCCAGAGCGCGCGCGCGCGGTGGACCGCGCG 422

Qy 301 gagggcttcttgcagccacgagaagctacaccgcttcctaaacggctgactctagt 360
Db 423 GAGGGTTCTTGTTCAGCCACAGAGAGCTACACCGCTTCTTAAACGCGCTGATCCTAGT 482

Qy 361 gtccagatcgccggaattttgctccggtgaatgaacagccgctccggcgttaatttcog 420
Db 483 GTTCAGATCGCCGGAATTTTGTCTCCGGTGAATGAACAGCCCGCTCCGGCGTAATCTTCG 542

Qy 421 gtggctcgaaactcccgattccatcaagagagtgatgtgcgaacgagactaaccca 480
Db 543 GTGGTCGGAANAATTTCCCGATTTCATCAAGAGAGTGTATGTGCGCAACGGAGCTAACCCA 602

Qy 481 cttcacgagccggtgacaggtcaccactcttcgacgagagcggatggtttcacgccgc 540
Db 603 CTTACAGACCCGTTGACAGTCACTTCTTCGACGGAGACGGTATGTTTTCACCCCGTC 662

Qy 541 aaattcgaacacggttcagctagctacgcttccgggtttactcagactaacgggtttgt 600
Db 663 AAATTCGAACCGGTTACGTAAGTACGCTTCCCGTTTACTCAGACTAACCGGTTTGT 722

Qy 601 caggaaactcaattgggtcgacgggttttccccaaagccatcggtagcttcacggccac 660
Db 723 CAGGAACGTCAAATGGGTGCACCGGTTTTCCTCCAAAGCCATCGGTGAGCTTTCACGGCCAC 782

Qy 661 accggtattgcccactcatgtattctacgccagagctgcagccggtatagtcacccg 720
Db 783 ACCGTTATTGCCGACTCATGTATTCTACGCCAGAGCTGCAGCGGTATATGTCACCCG 842

Qy 721 gcacaggaacccggtgtagtaaacccggttttggtctatttcaatggcggtattggct 780
```

RESULT 5



```
/codon_start=1
/evidence=not experimental
/product="F3f9.7"
/protein_id="AAF71811.1"
/db_xref="GI:8052547"
/translation="MGKLGKRVVVEERYTKPOGLYVNVKDVVKLKLKLVESKLAP
CYPDDESDHLEPCIFLYPLNSLRCCMKSCICTECFLOKMPNSARPTRQPLIN
LMGFFPVSLLFMKPCPKTNYAYEYGVKSKKEEGIEQVEEQRVLEAKIRMQKE
MODDERKORLESSESTSAEMTGYSTGFLVFWFAAIYSNLSMDGDEAPSQ
NASVYORSRREGRNEDVDLELMYMEALWLSVQETGTQRNSASGEITSSRYVT
DNHSYVSPRPVPTPIVEATPSSSGSLSCALSALEAOMYGESSHNNHNNVSSY
SMLPNCDSYVDIEQVDGIDNHHHRHHVEMGETSSNSYSSVYMTGEGHNPFP
PLVIYVPESEFQEMAMAVSAEVHATTTCAPTEVTQ"
CDS
complement(join(2808..23063,23118..23244,23511..23706))
/feature="hypothetical protein"
/codon_start=1
/evidence=not experimental
/product="F3f9.8"
/protein_id="AAF71810.1"
/db_xref="GI:8052546"
/translation="MSGRGKVKSEPMKVVFINTQYVETDARSFKTVVQELTGKNAIVA
AGPDSPAFDRCYDGGSKIGPITNVVFWKIMSYLSVRPMSNPKRAVYSPHMLFK
RTMKPTNYONKRCBCVTKALKATSLRSQASTMGFLWGVFATQVQLGITFFSIFEI
VLYFWFFSCKIMELTKILISLVLSKTV"
CDS
join(26752..26907,27020..27871,27923..28153)
/feature="similar to exopolysaccharuronase precursor
sp|Q00293|POLX_ASPTU"
/codon_start=1
/evidence=not experimental
/product="F3f9.9"
/protein_id="AAF71796.1"
/db_xref="GI:8052532"
/translation="MVSTISGFLFIAAVASSISAAPSAAALVGRKVFDRSYGARG
DKTNDTMAFTKAMQKGLPRVYIPFGTFYLGAVAFGPCKSRISFTIKGTLLA
PKDPAIKQDSMIIPRYVDYLVSGGILDGGSYSWPLNNCRTHNCRALPMMWFO
FVRSRLRIKISINEMGHLPFFSVQHEFDITRVNIKAPGDSPTDGIKIGSNHMKIF
HVDIATGDDCIALLSGTNLDINKVCGFHGHSVSLGKFKGKSVGLLVIRNSIFN
GTSNGVRITKPSPEPLNVSNFLPNLQMDVQSPINIDQRYCPNPSPQVTSLTR
NKFSKIQIRVFKFQNIWGTSTAKEAVLQCSKNVPCKNVQLFNINIYHRGRDGPATS
VCENVGWGGIKSPSPSCIR"
CDS
33566..35539
/feature="similar to 9-cis-epoxycarotenoid dioxygenase
gp|AAF26356.1"
/codon_start=1
/evidence=not experimental
/product="F3f9.10"
/protein_id="AAF71797.1"
/db_xref="GI:8052533"
/translation="MTLITISGMVYISLLSDAHHSQYQNTNLVLKPKIPQTA
FNOESTMASTLLPSTQFLDRFTSTSSSRPKLQSLFSSTLRNKKLVVPCYVS
SVNKKSSVSSLSQSPFTKPPSWKLCNDVTNLIKPTTNNQKLPVORVTAAMVLD
NAMIHSRRRHPKPTADPAQVIAAGNPPVPEKPVHNLVPTGTVEPCIOGVYVNGA
NMLHPKVSGLHLFDGDCGMHVARFEDNGSVYACRFTENRLVQRECRPVPPKAIGE
LHGLGIKMLFNLTRGLVDPPTGGIGLVANAGLVYFNGHLLAMSEDDLPYHVKTQ
TGDLETSRYDFDGLQKSTMIAPKPIDPETRELFALSYDVVSKPYLYKFRFTSDGKS"
Query Match 47.9%; Score 862.8; DB 8; Length 95769;
Best Local Similarity 73.1%; Pred. No. 2.1e-245;
Matches 1139; Conservative 0; Mismatches 407; Indels 12; Gaps 2;
QY 245 acactaaacagatgaattgttcagagagcgcgcgagcgatgttgagcgcgaggg 304
Db 33987 AGAATCCAAAATTAACCCCGTGCAGAAACCGCGCTATGCTTTTAGACGCGGTGAAA 34046
QY 305 gtttctgtgagcacag---aagctacacccgtctccttaaaacgctgactagt 361
Db 34047 AGCCATGATCTCACACGAGCGCGCGTATCCCCCATCTTAAACCGCGGATCTCGCG 34106
QY 362 ttcagatcccggaatttgcctcggtgaatgaacagccgctccggtgaattcttcgg 421
Db 34107 TCAATAACCGGGAACTTCTTCCCGGTACCGAGAAACCTGCTGCTACATACCTTCGG 34166
QY 422 tggctcgaaacttcccattccattcaaaaggagtgatgtgcgaacggagtaaccac 481
```

```
Db 34167 TGACTGGAACAGTACCAAGATGCATTCAGAGGAGTTTTCGTCAGAAACGAGCAATCCAC 34226
QY 482 ttcacgagccggtgacaggtccaccacttcttcgacggagagcggtatggttcacgcgtca 541
Db 34227 TTCACAAACCACTCTCCGGCCACCACTTTGTTGCGAGGAGCGGTATGGTACAGCTGTCC 34286
QY 542 aattcgaacacgggttcagctagctacgtctgctgcttactcagactaaacgggtttgttc 601
Db 34287 GGTTCGATAACCGCTCGGTAGCTAGCTAGCTGTCGGTTCACCGAAACAAACCGGTGGTTC 34346
QY 602 aggaacgccaattgggtcgaccgggtttcccccagaagccatcggtgagcttcacggccaca 661
Db 34347 AAGAACGTAATGTGTCGTGCTCGCGTTTCCCAAGAGCAATTCGAGAGCTTCACGGACAT 34406
QY 662 ccggtattgcccgcactcatgctattctacgccagagctgcagccggtatagtcgacccgg 721
Db 34407 TAGGAATGCCCAAGCTTATGCTTCTCAATACCGCGGGCTATTTGGGTTAGTCGACCCGA 34466
QY 722 cacaggaacgggtgtagtaaacgcgggtttggtctatttcaatggcgggtattgggcta 781
Db 34467 CCGGAGGACTCGGTGTCGCTAACCGCGGTTTAGTTTATTTCAATGGTCATCTCTTAGCCA 34526
QY 782 tctcggagagatgatttaccttaccaggttcagatcactcccaatggagatttataaacgg 841
Db 34527 TGTCCGAAGACGATTTACCGTACCATGTCAAAGTTACTCAAACCGGAGATTTAGAAACTT 34586
QY 842 ttggtcgggttcgatttgcgtgacgaattagataccacgaattgattccaccacgaagtcg 901
Db 34587 CGGTCGGTACGATTTTCAGCGTCACTTAAATCAACAAATGATAGCCCAACCGAAATCG 34646
QY 902 acccggaatccgggtgaactcttcgctttaagctacgagctgcttcaagcgccttaccctaa 961
Db 34647 ATCCGGAACCCGAGAACTTTCGCTTTTAAGCTACGACGTCGCTTTCAAAGCCTTACTTAA 34706
QY 962 ataacttcgattctcacggagcggaaataataaccggagcgtcagagctcagacttgatc 1021
Db 34707 AATACTTCAGATTACATCCGACGGTGAGAAATCAACAGACGTCGAGATTCGCTTGATC 34766
QY 1022 agccaacgagatgcacgatttcgagattacagaaacttcgctcgtcgtcactcgaccagc 1081
Db 34767 AACCAACGATGATCCCGATTTTCGGGATCACTGAGAAATTCGTAGTGTTCGAGACCAAC 34826
QY 1082 aagtcggttttcaagctgcggagatgcgcgggtgggtggttcctcgggtgggtttacgacaaga 1141
Db 34827 AAGTGTATTCGATTCCGATTCCGGAGATGATAAGAGGTGGTTCCTCCGCTGGTTTACGACGAGA 34886
QY 1142 acaaggtcgcaagattcgggatttttagacaataacgcgcaagattcctcgaacattaagt 1201
Db 34887 AGAAGAAATCAAGATTTCGGGATTTTGAATAAAACCGCAAGAGATGCTTCGTCGATTCAAT 34946
QY 1202 ggattgactccagattgcttctccttccatctcgaagcgttgggaagagcagagaa 1261
Db 34947 GGATCGAAGTACCAGATTGTTTTTTTTCATCTATGAAACTCTTGGGAAGAACAGAGAA 35006
QY 1262 cagatgaagtcgctgagatgggtcctgatgactccaccagactcaattttcaacgagat 1321
Db 35007 CAGACGAGTTTGTGCTGATGGATCATGATCAGCCACCTGATTCAATTTTCAACGACAC 35066
QY 1322 ctgacgagaatctcaagagtgctcctgtctgaatccgctcctgaatctcaaacgggtgaat 1381
Db 35067 ACGACGAACACTTCAGAGTGTGTTTGTCCGAGATAAGAGCTAAACCTTAAACAGGGGAT 35126
QY 1382 caactcgcgctccgactcatctccaacgaagatcaacagtcacacgtcaacgtcgaagcaggaagg 1441
Db 35127 CAACACGTAGACCGGTTATCTC-----GGAAACAGTTAATCTCGAAGCTGTGATGG 35177
QY 1442 tcaacagaaacatgctcgccgctaaaaacaaattcgttacttggcttttagccggagcgt 1501
Db 35178 TCAACCGAAATTTATTAGGTAGAAAAACCGGTATGCTTATCTGGCTTTTAAACCGAACCGT 35237
QY 1502 ggcctaaagtcaggatttcgctaaagttgactcactcactggagaagtttaagaacatc 1561
```

Db 35238 GGCCTAAAGTGCCTGTTTCGCTAAAGTGGACCTTATCGACCGGAGAGATTCCGAAAGTATA 35297  
Qy 1562 ttacggcgataaacggttacggagagagagcctctcttctcccccgaagagagagg 1621  
Db 35298 TTTACGGAGAGGGGAAATACGGAGGAGAGCCTCTCTTCTTACCTTCGCGTACGGAGAG 35357  
Qy 1622 aagcgaagagatcacatcctctgttccgttcacgacgagagaacatggaaatcggagttac 1681  
Db 35358 AAGACGGAGGTTTACATAATAGTGTTCGTTTCACGACGAGAGAGAGTGAATTCGGAACCTCC 35417  
Qy 1682 agatagttaaacccgcttagcttagaggttgaagcaacggttaaaccttccgtcaagggttc 1741  
Db 35418 AACTATAACGCCGCTTAATATGAAGCTTGAAGCTTACCGTAAACGCTTCGCTCGAGAGTGC 35477  
Qy 1742 cgtacggattttacgggttacatcctcgaagccgatgatttggcgaagcaggtcgtgtg 1799  
Db 35478 CTTACCGGTTTCACGGAACGTTTCATCAGTAAGGAAGATTATCGAAGCAAGCTTTGTG 35535

RESULT 6  
AX148316 AX148316 1818 bp DNA linear PAT 08-JUN-2001  
LOCUS  
DEFINITION Sequence 15 from Patent EP1116794.  
ACCESSION AX148316  
VERSION AX148316.1 GI:14347203  
KEYWORDS  
SOURCE tomato.  
ORGANISM Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;  
Lycopersicon.  
REFERENCE 1 (bases 1 to 1818)  
AUTHORS Iuchi,S., Kobayashi,M. and Shinozaki,K.  
TITLE Transgenic plants carrying neoxanthin cleavage enzyme gene  
JOURNAL Patent: EP 1116794-A 15 18-JUL-2001;  
Riken (JP)  
FEATURES  
source  
1..1818 Location/Qualifiers  
/organism="Lycopersicon esculentum"  
/db\_xref="taxon:4081"  
1..1818  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC41199.1"  
/db\_xref="GI:14347204"  
/translation="MAATTSHTATNTWIKLSMPSEKFGFASNSISLLKNQHNRSQSL  
NINGSLOAPPLIHFPKOSSNYOTPKNNTISHPKOENNSSSSSTSKWNLVQKAAMAL  
DAVSALTKEHLEHLPKTDPRVQISGNFAPVPENVCOSLPVTGKI PKCVQGVYVR  
NGANPLEPTAGHHFFDGDGMVHAVQPKNGSASACRPFETERLVQEKALGRPVFPKA  
IGELHSGIARLMFLFARGULFGLVDHSGTGVANAGLVFNRLLSBEDDLPHVK  
VTPTGDLKTEGRFPDQGLKSTMTIAHPKLDVPSGELFALSDYVIQKPYLYKFRFSKNG  
EKSNDVIPBDDPTMHDFAITENFVPIQQVYFKMSMI RGSPVYVSI DKNKYSRFG  
ILDKVADGSLDKWVEVDCPCFHLNWAWEAEDEIWI GSCMTPDPSI FNECDEGL  
KSVLEIRLNLTKCKSRKSIENPDQVNL EAGMVRNKRKTEYAYLAIAPWPWK  
VSFAKNLFTGEVEKFTIENDKYGGGEPLFLPRDPNSKEEDDGYLLAFVHDEKENKSE  
LQIVNMSLKLEATVLPFSRVPYFGHGFIFINANDLANQA"  
BASE COUNT 569 a 383 c 369 g 497 t  
ORIGIN

Query Match 45.8%; Score 824.6; DB 6; Length 1818;  
Best Local Similarity 69.6%; Pred. No. 2.5e-248;  
Matches 1134; Conservative 0; Mismatches 489; Indels 6; Gaps 1;  
Qy 170 cccagctcttcattccctaagcaatcatcaaacctctccgcccattgttgaagccca 229  
Db 185 CTTCAAAATTATCAACACCAAGAATAATACAAATTCACACCCAAACCAAGAAACAACA 244  
Qy 230 aagccaaagaatcccaacataacagatgaattttgttccagagcggcgagcggt 289  
Db 245 ACTCCTCTCTCTTCACTTCCAAAGTGGAAATTTAGTGCAGAAAGCAGCAGCAATGGCTT 304



```

Db 1385 AAACAGGGAATCAACAGAAATCATATCGAAACCCGGATGAACAGTGAATTTAG 1444
Qy 1430 aagcaggatggtcaacagaaacatgctcgccgtaaaacccaattcgcttacttgctt 1489
Db 1445 AAGCTGGAATGGTGAACCGAAACAAACTCGGAAGGAAACACAGATATGCTTATTTGGCTA 1504
Qy 1490 tagccagcggctgacctaaagctcaggttcgctaaagttagtcatcactatcgagaag 1549
Db 1505 TCGCTGAACCATGGCCCAAAAGTTCTGGTGTTCGAAAAGTAAACCTGTTTCAACCGGTGAAG 1564
Qy 1550 taaagaaacatcttacggcgataaccgttacggagagagcctctgtttcccgagag 1609
Db 1565 TTGAGAAATTCATTTATGTTGACAAATAATGTTGGGAACCTCTTTTACCAAGAG 1624
Qy 1610 -----aaggagagaggaagcgaaggatacatcctctgtttcgttcacgcagagaaga 1663
Db 1625 ACCCAACAGCAAGGAAGAGCATGTTATATTTTAGCTTTTCGTTACGATGAGAAAG 1684
Qy 1664 catggaatcggagttacagatagttaaacgcggttagcttagagttgaagcaacggtta 1723
Db 1685 AATGGAATCAGAACTGCAAAATGTTTAAACGCAATGAGTTTGAAGTTGGAGGCAACTGTGA 1744
Qy 1724 aacttcgcaagggttcggtacgatttcacggatcattcatcggagcggatgattgg 1783
Db 1745 AGCTTCATCAAGAGTCTTATGAGATTTCATGGAACATTCATAACGCCAATGATTGG 1804
Qy 1784 cgaagcagg 1792
Db 1805 CAATCAGG 1813

RESULT 7
LENEOXANT
LOCUS LENEOXANT 2171 bp mRNA linear PLN 22-MAY-1998
DEFINITION Lycopersicon esculentum mRNA for nine-cis-epoxycarotenoid
dioxigenase.
ACCESSION 297215
VERSION 297215.1 GI:2769641
KEYWORDS nine-cis-epoxycarotenoid dioxigenase.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 2171)
AUTHORS Burbridge,A., Grieve,T.M., Jackson,A., Thompson,A. and Taylor,I.B.
TITLE Structure and expression of a cDNA encoding a putative neoxanthin
cleavage enzyme (NCE) isolated from a wilt-related tomato
(Lycopersicon esculentum Mill.) library
JOURNAL J. Exp. Bot. 47, 2111-2112 (1997)
REFERENCE 2 (bases 1 to 2171)
AUTHORS Burbridge,A.
TITLE Direct Submission
JOURNAL Submitted (01-JUL-1997) Burbridge A., The University of Nottingham,
Physiology and Environmental Science, Sutton Bonington Campus,
Loughborough, Leicestershire, LE12 5RD, UK
REMARK Revised by [3]
REFERENCE 3 (bases 1 to 2171)
AUTHORS Burbridge,A.
TITLE Direct Submission
JOURNAL Submitted (09-JAN-1998) Burbridge A., The University of Nottingham,
Physiology and Environmental Science, Sutton Bonington Campus,
Loughborough, Leicestershire, LE12 5RD, UK
COMMENT On Jan 13, 1998 this sequence version replaced gi:2243153.
FEATURES
Location/Qualifiers
1..2171
/organism="Lycopersicon esculentum"
/db_xref="taxon:4081"
79..1896
/codon_start=1
/product="nine-cis-epoxycarotenoid dioxigenase"
CDS

```

```

/protein_id="CAB10168.1"
/db_xref="GI:2769642"
/db_xref="SPTREMBL:O24023"
/translation="MATTTSHTATNTWKTLSPSSKEFGFASNSISLLKQHNQSL
NINSSLOAPPTLHPKOSNNVTPKNTTISHPKOENNNSSSSSTSKNVLVKAAMAL
DAVESALTKELEHPLPKTADPRVQISGNFAPVPENPVCSLPVTGKTPKCVQGVYVR
NGANPLFPTAGHHFFDQGMVHAFQKNGSASYACRPTETLRLVQEKALGRVPVPKA
IGELHSGIARLMFLFYARGLFGVDSKGTGVANAGLVYFNNRLLAMSEDDLPYHVK
VTPGDLKTEGRFDQGLKSTMTAHPLKDPVSGELFALSVDVIQKPYLTKFRFSKNG
EKSNDVEIPVEDPTMHDFAITENFVVPDOQVVKMSMIRGGSPVVDKNKVSFPG
ILDKYAKGSDLKWVEYDCPCFHLNMAWEAEDEIVIGSCMTPPDSINEDEGL
KSVLSELRNLTKSTKRSIIENPDQVNLKAGVNRNKLGRTEYAYLAIAEPWPK
VSGPAKYNLFTGEVKFYGGGFLFPLRDPNSKEEDDGYILAFVHDEKEWKSE
LQIVNAMSLEATVKLPSPVYFGHGFNFNANDLANQA"
BASE COUNT 688 a 435 c 431 g 617 t
ORIGIN

Query Match 45.8%; Score 824.6; DB 8; Length 2171;
Best Local Similarity 69.6%; Pred. No. 2.6e-234;
Matches 1134; Conservative 0; Mismatches 489; Indels 6; Gaps 1;

Qy 170 ctceagctctcatttccttaagaatcatcaaacctctccgccttctgttaagccca 229
Db 263 CTTCAAAATTAACAACACCAAGAATAATACAAATTCACACCCAAAACAAGAAAAACA 322
Qy 230 aagccaaagaatcccaactaaacagatgaattgttccagagagcggcggcagcgt 289
Db 323 ACTCCTCTCTTCTCAACTTCCAAGTGAATTTAGTCGAGAAACAGCAGCAATGGCTT 382
Qy 290 tggacgcggcggagggttccctgtcagccacgagaagctacacccgcttctcctaaacgg 349
Db 383 TAGATGCTGTAGAAAGTCTTTAACTAAACATGAACCTTTGAAACACCTTTTGCAGAAAACAG 442
Qy 350 ctgacctagctcagatccgcggaatttctcctcggatgaataacagccgcgcggc 409
Db 443 CGACCCACGAGTCCAGATTTCTGGGAATTTTGTCTCCGGTACCGGAAATCCAGTCTGTCT 502
Qy 410 gtaactctccgggtggtcggaactcccgatccatccatcaaggagtgatgtgcgcaacg 469
Db 503 AATCTCTTCGGGTACCCGGAAAAATACCAATGTGTTCAAGCGCTTTACGTTCCAAACG 562
Qy 470 gagctaacccacttcacagcgggtgacaggtcaccaactcttcacgagagcgtatgg 529
Db 563 GAGCTTAACCCCTCTTTTGAACCAACCCGCGACACCAATTTCTTCACGCGCAGCGTATGG 622
Qy 530 ttcaacgcgtcaaatcgaacacggttcagtagctagctgacggttgcgggttactcagacta 589
Db 623 TTCACGCGCTCAATTTCAAAAATGGGTGCGGTAGTTACGCTTGCCTTTTCACTGAAACAG 682
Qy 590 accggtttggtcaggaacgctcaattgggtcgacgggttttccccaaagccatcggtgagc 649
Db 683 AGAGCTTGTTCAGAAAAAAGCTTGGGTGCGCCCTGTTTTCCCTAAAGCCATTTGTTGAAT 742
Qy 650 ttacgcccacaccgggtattgcccgaactcatgtattctacgcccagagctgagccggta 709
Db 743 TACATGGTCACTCTGGAATTCGAAGCTTATGCTCTGTTTACGCTCGTGGGCTCTTTCGAC 802
Qy 710 tagtcaccccgccacacggaacccggttagttaaaccggttttggtctatttcaatggcc 769
Db 803 TTGTTGATCACAGTAAGGAACTGTGTGTGCAAAACGCGCGTTTAGTCTTATTTCAATAACC 862
Qy 770 ggttattggctatgcggagatgatttaccttaccaggttcagatcactcccaatgag 829
Db 863 GATTACTTGTCTATGCTCGAAGATGATTTGCCCTTACCATGTAAAGGTAAACCCACCGCGG 922
Qy 830 atttaaaacccgttggtcgggttcgattttgatggacaattagaaatcccaatgattgcc 889
Db 923 ATCTTAAACAGAGGCTCGATTTCGATTTTCGACGGCCAGCTAAATCCACCATGATAGCTC 982
Qy 890 accgaaagtcgacccggaatcccggtgaactcttcgctttaaagtaacgacgtcgtttcaa 949
Db 983 ACCCAAGCTCGACCCAGTTTCCGGTGAGCTATTTCGCTTCTTAGCTACGATGTGATTCAGA 1042

```





Db	758	GGCTTATGCTGTTTTACGCTCGTGGGCTCTTCGGACTTATTGATCACAGTACAGAACTG	817
Qy	734	gtgtagctaacgcgcggttttggtctatttcaatggccggttatgtgctatgtcggagatg	793
Db	818	GAGTTGCAAAAGCCGGTTTGTAGTCTATTTCAATAACCGATTACTTGTCTATGTCTGAAGATG	877
Qy	794	atttaccttaccagtttcagatcactcccaatggagatttataaacogtltggttcggttcg	853
Db	878	ATTTGGCTTTACCATGTAAAGGTAAACACCCACCGGGCATCTTAAACACAGAGGTCGATTCG	937
Qy	854	attttgatgacaatttagaatcccaatgattgccaccgccgaagtgcgaccggaatccg	913
Db	938	ATTTTCGACGGCAGCTTAAATCACCATGATAGCTACCCAAAGCTCGACCACGTTTCCG	997
Qy	914	gtgaactcttcgctttaagctacgacgctcggtttcaaaagccttacctaaatctactccgat	973
Db	998	GTGAGCTATTGTCTTTAGCTAGCTGATTCAGACGCCATACCTCAAGTATTTTCAGAT	1057
Qy	974	tctcaccggacgaactaaatcacccggacgctcgagattcagcttgatcagcacaacgatga	1033
Db	1058	TTTCAAAAATTCGGGAAAAATCAATGATGTTGAAATCCAGTTGGAAGACCCCAACAATGA	1117
Qy	1034	tgcacgatttcggtattacagagaacttcgctcgttcgctaccctgaccagcaagctcgtttcca	1093
Db	1118	TGCATGATTTTCGCAATTTACTTGAATAATTCGTGATCATTTCTCTGATCAACAAGTCGTTTTCA	1177
Qy	1094	agctcgccgagatgatccgcggtgggtctccggtgggtttacgaacaagaacaaagtcgcaa	1153
Db	1178	AGATGCTGAAATGATCCGTGGAGGTTTACCGGTGGTTTTACGAAGAACAANAAGTGTCCC	1237
Qy	1154	gattcgggatttttagacaaatccgcgaagattcatcgaaactaaagtggaattgatgctc	1213
Db	1238	GATTTGGGATTCTGGATAAGTACCGCAAAAGATGGGTCTGATTTTGAAATGGGTTCAGTTC	1297
Qy	1214	cagattgcttctcgtctccatctcttgaaacgcttgggaagagccgagaacagatgaagtcg	1273
Db	1298	CTGATGCTGTTCTGTTTCCATCTCTGGAATGCTTGGGAAGAACCAGAAACAGATGAAATCG	1357
Qy	1274	tcgtgatagggtcctgatgactccaccagactcaattttcaacgagctcgaagagatc	1333
Db	1358	TTGTATTTGGTTTCATGTATGACACCAACCAAGATCTCAATTTCAATGAATGTGATGAAGAC	1417
Qy	1334	tcaagagtgctcgtctgaaatccgcctgaattctcaaacccggtgaatcaactccgcgtc	1393
Db	1418	TAAAAAGCGTTTTATCCGAAATTCGTCTCAATTTTGAAACAGGGAAATCAACAAGAAAG	1477
Qy	1394	cgatcatctccaacgaagtcaacaagtcacacctcgaagcgggagtggtcaacagaaca	1453
Db	1478	CCATAATCGAAAAACCCGGATGAACAGTGAATTTAGAGGCTCGAATTGTGAACAGAAACA	1537
Qy	1454	tgctcggcgtaaaaacaaatcgtttacttgcgtttagccgaccgtgcccataaagtct	1513
Db	1538	AAC TTGGAAGAAAAACACATGATGCTTTATTTGGCTATTGCTCGAACCAATGGCCAAAGTTT	1597
Qy	1514	caggattcgcgtaaagttgatcttcaactactgggaagtttaagaaacatctttacggcgata	1573
Db	1598	CTGGTTTGGAAAAGTAGACCTGTTTACAGGTGAAGTTGAGANAATTCATTTATGGTGACA	1657
Qy	1574	acggttacggaggagagcctctgtttctccccggag-----aaggaggagaggaagcgc	1627
Db	1658	ACAAATATGTTGGGGAGCCTCTTTTTTTTACCAAGAGACCCAAAACAGCAAGGAGGAACGC	1717
Qy	1628	aaggatacatcctctgttttcggttcacgacgagaagacatggaaatcggagttacagatag	1687
Db	1718	ATGGTTATATTTTATGCTTTCGTTTCATGATGAGAAAAGATGGACATACAGAGCTTCAAAATG	1777
Qy	1688	ttaacgcgcttagcttagaggttgaagcaacggtttaaacttccgtcaagggttccgtacg	1747
Db	1778	TTAATGCAATGACTTTAAAGTTAGGCCAACAGTGAAACTTTCCATCAAGAGTTCCTTATG	1837
Qy	1748	gatttcacggtacattcatcggagccgatgatttggcgaagcagg	1792

Qy 593 gggttggttcaggaaacgtcaattgggtcgacgggttttccccaaagccatcggtgagcttc 652  
Db 641 GTCCTCGCAGAGAGAAATCTTAGCGCGCCCGTGTTCCTCCGAAGCCATCGGGAGCTCC 700  
Qy 653 acggccacacccggattatgccegaactcatgctattctacgcccagagctcgacggcgatag 712  
Db 701 ACGGCACTCGGGCATCGCGGCTCCTCTCTTCTACGCGCGGTCTCTTCGGGCTCG 760  
Qy 713 tccagccggcgcacacggaacccggtgtagctaaacgcccgggtttgggtctatttcaatggccgg 772  
Db 761 TTGATGGGTCCAGGGCATGGCGGTGGGAACGCGCGTCTCGTCTACTTCAACACCAACC 820  
Qy 773 tctggctatgctcgagagatgattacattaccagagttcacatgctcactcccaatggagatt 832  
Db 821 TCTTGCCATGTCCGAAGACGATTTACCTTACCAGTGAAGATCACCTTAAACGGCGACT 880  
Qy 833 taaaaacggttggtcggttcgatttgatggacaattagaaatccacaatgatgtgccacc 892  
Db 881 TAAACACCGTTGGCGTTAGCACTTCAACGGGCAGCTCAACTCAACAATGATCGGCCACC 940  
Qy 893 cgaagtcgacccggaatccggtgaactcttcgcttttaagctaacagctcggtttcaagc 952  
Db 941 CGAAACTGACCCCTCGACGCGGACCTCCACGCGCTCAGCTACGACGTCAATTCAGAAGC 1000  
Qy 953 ctactctaaataacttcgattctcaccgacggaactaaataaccgacgctcgagattc 1012  
Db 1001 CTTACTCTAAGTACTTCGTTTCTCCCGCGCGGTCAAGTCCCCCGACGTGGAATCC 1060  
Qy 1013 agcttgatcagccaacgatgatgcacgatttcgagatttacagagaacttcgctcgctac 1072  
Db 1061 CCCTGAAGAGCCACCATGATGCAGATTTCCGCATTAACGGAGAAATTCGTGCTGCTC 1120  
Qy 1073 ctgaccagcaagtcggttttcaagtcgcccggagatgatccgcccgttggtttctcgggtg 1132  
Db 1121 CCGACGACGAGTGTGTTCTTCAAACTAACGGAGATGATCACCGCGCGTCCCCCGGTG 1180  
Qy 1133 acgacaagaacaggtcgcaagattcgggatttttagacaataaccgcaagattcatcga 1192  
Db 1181 ACGACAAGAACAAAACCTCACCGTTTGGGATTTCTGCAGAAGAAATCGAAGGACGGAATG 1240  
Qy 1193 acatttaagtggattgatgctccagatgctctgcttccatctctggaacgctttgggaag 1252  
Db 1241 CGATCGGTGGATGACGCGCGCGGATGTTTCTGCTTCCACCTCTGGAACGCGTGGAGG 1300  
Qy 1253 agccgaaacagatgaagtctcgtagatgggtctctgtagtactccacagactcaattt 1312  
Db 1301 AGCCCGAAACCGAGGAGTGTGGTGAATGGGTCTGTCATGACCCCTCGCGACTCCATTT 1360  
Qy 1313 tcaacgagctgacagagaatctcaagagtgctctgctgtaaatccgctgaatctcaaaa 1372  
Db 1361 TCACGAATGCCAGGAGAGTTTGAAGAGCTGCTGTTCAGAGATAGGCTGAACCTTGAGGA 1420  
Qy 1373 csggtgaatcaactcgcgctccgatcatctccaaacgaagatcaacaagtcaacctcgaag 1432  
Db 1421 CGGCAAGTCCATCGGCGCCCAATATCTCCGACGCCG---AACAAGTGAACCTGGAG 1477  
Qy 1433 caggatggtcaacagaaacatgctcgccgctaaacaaatctgcttacttgggttttag 1492  
Db 1478 CCGCATGTTGAACAGAAACAAGCTCGGAAGGAAGACCCAGTTCGCGTATCTGGCTCTCG 1537  
Qy 1493 ccgagcgtggcctaaagctcaggattcgctaaagttgatctcactcaggagaagtta 1552  
Db 1538 CGAGCCCTGGCCCCAAAGTCTCGGCTTTGGCAAGTTGATTTGCTGAGTGGGGAAGTGA 1597  
Qy 1553 agaaacatctttacggcgataaccgttacaggaggagagcctctgtttctccccggagag 1612  
Db 1598 AGAAGTACATGATGAGAGAGAGAGTTCGTTGGGGAGGCTCTGTCTTCTCC---AAGC 1654  
Qy 1613 gaggaggaagacgaagatatactctgtttcgttcaacagcagagagacatggaat 1672  
Db 1655 GCCAAAAGAAAGACGATGGGTATTTCTGGCATTCGTGCACGACGAGAAAGAAATGGAAT 1714  
Qy 1673 cggagttacagatagttaacgcccgttagcttagaggttgaagcaacggttaaacctccgt 1732

Db 1715 CCGAGCTGCAGATTGTGAATGCCCAAAATTTAAAGCTCGAAGCTTCATCAAACTCCCT 1774  
Qy 1733 caagggttcgtagcagatttcacaggtacattcatcgaagccgatgattggcgaagcagg 1792  
Db 1775 CTCGTGTCCTACGGTTTTTCATGGAACCTTTCATTTCATTCACAGGATTTGAGGAACAAG 1834  
RESULT 10  
AB030293 Vigna unguiculata CPRD65 mRNA for neoxanthin cleavage enzyme, complete cds. 2432 bp mRNA linear PLN 01-DEC-2000  
DEFINITION  
AB030293 Vigna unguiculata CPRD65 mRNA for neoxanthin cleavage enzyme, complete cds.  
ACCESSION  
AB030293  
VERSION  
AB030293.1 GI:9857289  
KEYWORDS  
neoxanthin cleavage enzyme.  
SOURCE  
Vigna unguiculata cDNA to mRNA.  
ORGANISM  
Vigna unguiculata  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Papilionoideae; Phaseoleae; Vigna.  
REFERENCE  
1 (sites)  
Iuchi,S., Kobayashi,M., Yamauchi-Shinozaki,K. and Shinozaki,K. A stress-inducible gene for 9-cis-epoxycarotenoid dioxygenase involved in abscisic acid biosynthesis under water stress in drought-tolerant cowpea  
Plant Physiol. 123 (2), 553-562 (2000)  
JOURNAL  
MEDLINE  
20317197  
REFERENCE  
2 (bases 1 to 2432)  
Iuchi,S.  
AUTHORS  
Direct Submission  
JOURNAL  
Submitted (16-JUL-1999) Satoshi Iuchi, RIKEN, Plant Mol. Bio.; 3-1-1 Koyadai, Tsukuba 305-0074, Japan  
(E-mail:iuchi@tc.riken.go.jp. Tel:81-298-36-4359)  
FEATURES  
Location/Qualifiers  
1..2432  
/organism="Vigna unguiculata"  
/db\_xref="taxon:3917"  
126..1964  
/genes="CPRD65"  
126..1964  
/genes="CPRD65"  
/codon\_start=1  
/product="neoxanthin cleavage enzyme"  
/protein\_id="BAB11932.1"  
/db\_xref="GI:9857290"  
/translation="MPSSASNTWFNATLSPFPKDLPSSTPTNLLPLPKRTSSNTIT CSLOTLHPKQVOTSTSTSTATTTPPTIKTTITTTTPRETNPISDNOGLPVQK NFKQAATALDAVETALVSHKHPKPTADPRVOIAGNFAPVPEHAADQGLPVQEK IPKIDGVYVRNGANFLYPERVAGHHFFDGDGMVHAKFTNGASAYACRFTTQRLSQE KSLGRPVFKAIGELHGHARLLFYARGLFGLVDGSGMGVANAGLYVFNHLLIA MSDDLFPYHVRITPNDLTITVGRYFNQNLSTMIAPKLDPDVDGLHALUSDYIQK YLKYAFSPDGKSPDVEIPLKEPTNMHDFATITENFVVPDQGVFKLTMTTGGSPV VYDNKTSRFGILHNAKADANARWIDAPDFCFHLNANAMEETEVVVGSMTPA DSIPLNCEBSLSKSLSEIRLNLTKGSTRPPIISDAEONVLEAGWNRNKLGRKTPA YLALAPFPKVSFGAKVDLLSGEVKKYMGEEKFGPEPLFLPNQKEDDGYILAFVHD EKSWKELQIVNAQNLKLEASIKLGRVPYGFHGTFTHSKDLRQA"  
BASE COUNT 654 a 711 c 533 g 534 t  
ORIGIN  
Query Match 44.4%; Score 798.4; DB 8; Length 2432;  
Best Local Similarity 70.4%; Pred. No. 1.7e-226;  
Matches 1098; Conservative 0; Mismatches 456; Indels 6; Gaps 2;  
Qy 233 ccaagaatccaacactaaacagatgaattgttccagagagcggcgcgagcgttg 292  
Db 406 CCAACCAACCAATACCTCAAAAATGGAACTTTCTCCAGAAAGCCGCTGCCACGCGCTTGG 465  
Qy 293 acgcgcgaggaggtttccttgcagccacagagaagctacacccgcttctctaaaaacgctg 352  
Db 466 ACCTGGTGAACCGGCTCGTCTCGCACGAGCGCAAAACACCCCGCTCCCAAAACGGCGG 525



QKPYLYFRSADGVKSPDVEIPLKEPTMMHDAIFATNFVVDQVVFVFLTBMITGG  
SPVYDKNKTSRFGILDKNAKDNAMRWIDAPECFHLNANWEEPTDRIIVIGSCM  
TPADSIENECDESLKSVLSERLNLRGKSTRPIISDAEQVNLNAGWVNLKGRKT  
QFAYLALAEWPVKVSGFAKVDLFSGEVQKTYMGEEKGEPFLFDPNGEEEDGTYLAF  
VHDBEKWSBLQIVNAQNLKLEASIKLPSRPVPGFHGTFIHSKDLRKQA"  
BASE COUNT 627 a 687 c 516 g 568 t  
ORIGIN

Query Match		44.0%;	Score 791.6;	DB 8;	Length 2398;
Best Local Similarity		70.0%;	Pred. No. 1.8e-224;		
Matches 1096;		Conservative 0;	Mismatches 464;	Indels 6;	Gaps 2;
Qy	227	ccaaagccaaagatcccaactaaacagatgaattgtccagagagcgcgcgag	286		
Db	407	CGACACCAACAACCATTTCTCNAAAATGAACTTTCTGCAGAAACCCGCTCGACAG	466		
Qy	287	cgttggacggcgaggggttctctgtcagccagagaagctacacccggttctctaaaa	346		
Db	467	GCTTGGATATGTTGGAGACGGCGTGGTTTCGCACGAGCAACACCCCGCTCCCAAAA	526		
Qy	347	cggtgatcctagtgttcagatcgccggaattttgtcccggtgaatgaagacccgtcc	406		
Db	527	CGCGGACCCGAAATGTCAAATCGCGGGAATTCGCGCGGTGCGGAGCAGCGCGCGG	586		
Qy	407	ggcgtaatcttcgggtggtcgaaaaactcccgattccatcaaaaggagtgtatgcca	466		
Db	597	ATCAAGCCCTTCGGGTGTCGAAAAATTCCTCCAAATGATTTGACGGCGGTGACGTGCGCA	646		
Qy	467	acggagctaacccacttcacgagcgggtgacaggtcacacttcttcgacggagacgga	526		
Db	647	ACGGCGCAATCCGCTTACGAGCCCGTCGCGGGCACCACTTCTTCGACGCGCAGCGCA	706		
Qy	527	tggttcacgcgctcaaatccgaacacggttcagctagtagtcggttgcgggttactcaga	586		
Db	707	TGTTCCACGCGCTGAAGTTCAGAACGCGCTGCGCAGTACGCGTGCCTTCACCGAAA	766		
Qy	597	ctaaccggttgttcagaaagctcaattgggtcgacgggttttcccaaaagccatcggtg	646		
Db	767	CACAGCTCTCGGCAAGAGAAATCTTTAGCGCGCCCGGTGTTCCCGAAGGCTATCGCGG	826		
Qy	647	agcttcacggccacacgggtattgcgactcatgctatttcagccagagctgagcgg	706		
Db	827	AATCCACGGCCACTCCGGCATAGTCCGCTCCCTCTTCTACGCTCGCAGCCTCTTCC	886		
Qy	707	gtatagtcgacccggcacacggaaccggtgtagctaaacgcccgggttggcttatctaatg	766		
Db	887	AGCTCGTTGATGGGTCCACGCGCATGGGGGTGGCGAACGCGGTCTCGTCTATTTCACA	946		
Qy	767	gccgggttattggctatgtcggagagatgatttaccttaccagagttcagatcactcccaatg	826		
Db	947	ACCACCTTTGGCCATGTCGGAAGACGATTTTACCATACACGATGAGATACCTCTAACG	1006		
Qy	827	gagatttataaacgggttgggtcggttcgatttggagacaattagaaatcccaaatgattg	886		
Db	1007	CGACTTAAACACCGTTGCGCGTTACGACTTCAAGCGCCAGCTCAACTCCCAATGATCG	1066		
Qy	887	cccacccgaaagtgcacccggaatcccggtgaaactcttcgctttaagctacagctcggtt	946		
Db	1067	CCACCCGAAACTCGACCCCGTCAACGCGGACCTCCACGCGCTCAGCTACGACGTCGTT	1126		
Qy	947	caaaagcttacctaaatcttcgattcttcacggacgggaactaaatcacggagcgtcg	1006		
Db	1127	AGAGCCCTTACTCTAAGTACTTTCGCTTTCTCCGCTGACGGCGTAAAGTCCCCCGACGTCG	1186		
Qy	1007	agattcagcttgatcagccaaagatgatgcagatttcgcatgattacagagaacttcgctcg	1066		
Db	1187	AAATCCCTCTGAAGAGCAACCATGATGCAGATTTTCGCAATAACGAGAAATTCGTCG	1246		
Qy	1067	tcgctacctgaccagaagctcgttttcaagctgcccggagatgatccggtgggtctccgg	1126		
Db	1247	TCGTCCTCCCGACGACGAGTGTGTTCTTCAAACTAACAGAGATGATCACCGCGGCTCCCCCG	1306		

Qy	1127	tggtttacgacaagaacaggtcgcaagattcgggatttttagacaaataaccccaagatt	1186		
Db	1307	TGGTCTACGACAAGAAACAAACCTCACGGTTTGGGATTTTGGATAAGAAATCGAAGGACG	1366		
Qy	1187	catcgaaacattaaatggattgagctccagatcgcttctcgttcctcattcctcgtggaacgctt	1246		
Db	1367	CGAATGCGATGCGGTGGATCGACCGCGCGGAGTGTCTTCTGCTCCATCTCTGGAACGGT	1426		
Qy	1247	gggaagacgacgaacacagatgaagtcgctgatagggtcctgtatgctcaccacagatt	1306		
Db	1427	GGGAGGAGCCTGAAACCCGACGAGATTGTGGTATTTGGGTCTTCATGACCCCTCGGACT	1486		
Qy	1307	caattttcaacgagctctgacgagaaatctcaagagtcctgtctgaaatccgctcgtgaatc	1366		
Db	1487	CCATTTTCAATGATGCGATGAGAGTTTGAAGACGCTTGTCTGAGATAAGGCTGAAT	1546		
Qy	1367	tcaaaacgggtgaatcaactcgccgtccgatcatctccaaagaagatcaacaagctcaacc	1426		
Db	1547	TGAGGACGGGCAAGTCGACTCGCCGCCCATTTATCTCCGACGCGG--AACAAAGTGAACC	1603		
Qy	1427	tcgaagcagggatggtcaacagaaaaatgctcgccgtaaaaccaaattcgtcttacttgg	1486		
Db	1604	TTGAAGCCGCGCATGCTGAACAGAAACAAAGCTAGGAAGAAACACTCAGTTCGCTTATCTGG	1663		
Qy	1487	ctttagccgagcggcgtggcctaaagtctcagagattcgctaaagtgtgactcactactggag	1546		
Db	1664	CTCTGGCGAGCCCTGCGCCCAAGTCTCGGGCTTTGCGAAAGTTGATTTGTTTCAGTGGGG	1723		
Qy	1547	aagttaagaacacatctttacggcgataaacggttacggagagagacccctctgttctcccg	1606		
Db	1724	AAGTGCAGAGTACATGTATGGAGAGAAAGTTTCGGTGGGGAGCCTCTGTCTTCTCCC-	1782		
Qy	1607	gagaagagagagagagagagagagatacatcctctgtttcgttcacagagagagacat	1666		
Db	1783	--AACGCGGAGAGAGAGAGTATGATTTCTGGCATTTCTGTCACGACGAGAGAAAGT	1840		
Qy	1667	ggaatcggagttacagatagttaaacggcttagcttagaggttgaagcaacgggttaaac	1726		
Db	1841	GGAATTCGAGCTCGAGATTTGTGAATGCCCAAAATTTAAAGCTCGAAGCTTCCATCAAC	1900		
Qy	1727	tccggtcaaggggttccgtracggatttcacgggtacattcatcggagcggatgattggcga	1786		
Db	1901	TCCCTCTCGGTTCTTATGTTTTCATGGCACTTTTATTCATTCTTAAGGATTGAGGA	1960		
Qy	1787	agcagg 1792			
Db	1961	AGCAAG 1966			

RESULT 12  
AC074176/c  
LOCUS AC074176 72058 bp DNA linear PLN 19-JAN-2001  
DEFINITION Arabidopsis thaliana chromosome 1 BAC T2H7 genomic sequence,  
complete sequence.  
AC074176  
VERSION AC074176.5 GI:12321625  
KEYWORDS HTG.  
SOURCE thale cress.  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 72058)  
Lin, X., Kaul, S., Town, C.D., Benito, M.-I., Creasy, T.H., Haas, B.J.,  
Wu, D., Maiti, R., Ronning, C.M., Koo, H., Fujii, C.Y., Utterback, T.R.,  
Barnstead, M.E., Bowman, C.L., White, O., Nierman, W.C. and Fraser, C.M.  
Arabidopsis thaliana chromosome 1 BAC T2H7 genomic sequence  
Unpublished  
2 (bases 1 to 72058)  
Town, C.D. and Kaul, S.  
Direct Submission  
TITLE Submitted (15-JUL-2000) The Institute for Genomic Research, 9712  
Medical Center Dr, Rockville, MD 20850, USA, cdtown@igr.org













exon /number=10  
complement(7566..7702)  
/gene="F28J12.20"  
/number=11  
exon 8557..8838  
/gene="F28J12.30"  
/number=1  
gene 8557..13947  
/gene="F28J12.30"  
join(8557..8838,8913..9050,9136..9198,9284..9373,  
9450..9572,9668..9755,10089..10163,10261..10280,  
10348..10368,12117..12725,12806..13321,13362..13622,  
13693..13947)  
/gene="F28J12.30"  
/note="similarity to MucD - Pseudomonas aeruginosa,  
PACHX:G1184684  
/codon\_start=1  
/contains EST gb:T22208"  
/product="putative protein"  
/protein\_id="CAA16717.1"  
/db\_xref="GI:2832642"  
/translation="MALASSKAFSSIFNTLSPINOSKFLVACSGSNHVDVDRRRIM  
IFGSSLLATSSLLGNSOORLPWESAIALAEQKEKEELEEEERENVALFKQTSFVY  
IEALELPTSSGDILTDEBNGKIEGTGSGFVWDKLGHIYVNHVIAKLDOFLQRC  
KSVLDVAKGTRFSKEGKIGLDPDNDLAVLKITEGRELNPVLGTSNDRVGOSCF  
IGNPYGENTLTIQVSGLGREIIPSPNGKISIEAIOTDADINSNGSGPLDYSYHTI  
GVNTATFTRKSGMSGLIETDSVLNSVYLTMVERKKQIQRNNSESNNRQKRIS  
HDKNRDELVYRILCPIDVVGVIKSGKVINAI RHNTAKI KVFQDLHGCSSRVIT  
IYCSVKRQKEIIGTKSENEPLCCAQDALLAVDAIVASDEENKTKNDRDDNKECR  
LLVFPSSSLIGKAGENIKRIRRTASVKVSKVSDPSHVCMEDVNVVLSGSP  
ESVKAULFAVSAIMYKNPRENIPLDTSQDVPAASVIVPSDLNSVYPTQGFYSND  
HILCOOGAPSPYFNALSDPQGAETAANPVPFASLPPVTHFGSGSSSEELVFKV  
LCPLCNMRVITKGGSTIKRIRREASGSCIEVNDRTKCGDDECVIIVTATEILFCCLS  
TPFVFMOPDDMKMAVEAVLLQEYINDEDAENKVMOLLVSKVIGVIGKSGSVIN  
EIRKTNANICISKGGKDDLVEVSEVSVFDALIOIVLRREDVLDGKDSVATRKPP  
ARTDNCFLSSSNAGYTLPSFMSMASTSGFHGFGFPPVKSCILDFH"  
8839..8912  
/gene="F28J12.30"  
/number=1  
exon 8913..9050  
/gene="F28J12.30"  
/number=2  
intron 9051..9135  
/gene="F28J12.30"  
/number=2  
exon 9136..9198  
/gene="F28J12.30"  
/number=3  
intron 9199..9283  
/gene="F28J12.30"  
/number=3  
exon 9284..9373  
/gene="F28J12.30"  
/number=4  
intron 9374..9449  
/gene="F28J12.30"  
/number=4  
exon 9450..9572  
/gene="F28J12.30"

Query Match 41.0%; Score 738.2; DB 8; Length 110102;  
Best Local Similarity 68.2%; Pred. No. 3.2e-208;  
Matches 1063; Conservative 0; Mismatches 478; Indels 18; Gaps 2;

Qy 244 aacactaaacagatgaattgttcagagagcggcgagcgttcgagcggcgag 303  
Db 2180 AATCCTCTCCGCTAAACATCTTCAGAAAGCGCGCGATTGCGATCGCGCTGAG 2239  
Qy 304 ggtttccctgtcagccagagaagctacacccgcttctctaaacggctgacctagtgt 363  
Db 2240 CGTGATTAATCTCACAGCAGCAAGATTCTCCATTCCTCAAAACCCTGATCCAGTGT 2299

Qy 364 cagatcgccggaatttttgcctccggtgaatgaacagcccgctccgcgtaattttccgggtg 423  
Db 2300 CAGATTGCGGGAATTAATCCCGGTACCGGAATCTTCGTCGCGGAACCTCACCGTC 2359  
Qy 424 gtcggaacttcgcgattccatcaagagtgatgtgcgaacggaggaactaacactt 483  
Db 2360 GAAGGAACAATCCCTGACTGCAATGACCGTGTATATCCGTAAACGGCGCAATCCGATG 2419  
Qy 484 cagagccggtgacaggtcacactcttcgacggagcgggtatggttcacgcccgtcaaa 543  
Db 2420 TTTGAGCCAAACAGCTGGGCACCAATTTATTCGACGAGACGGAATGTTTCACGAGTAA 2479  
Qy 544 ttcgaacacggttcagctagcttaagcttcgcccgtttatctcagactaacggtttgttcag 603  
Db 2480 ATAACCAACGGTTACGTAGCTACGATCGCGTGTACAAAACCGAGAGATTGGTTCAG 2539  
Qy 604 gaaagtcgaatgggtcgaccggttttcccaaacatcggtagcttcacggccacacc 663  
Db 2540 GAAAAACGATTGGGTGACACAGTTTTTCCGAAAGCAATCGGCGAGCTTCACCGTCACTCG 2599  
Qy 664 ggtattgcccgaactcatgctattctacgccagagctgcagccggtatagtcgaccggca 723  
Db 2600 GGAATCGCACGTTTGATCTGTTTTACGCACGTGGGCTTTGTGGTCTGATCAACAACCA 2659  
Qy 724 caggaacccggtgtagctaacgcccgttttgctctatttcaatggccggtatttggctatg 783  
Db 2660 AACCGCGTCGAGTAGCAAAACGCGGTTTGTGTTTACTTTAATAACCGGCTTTTAGCTATG 2719  
Qy 784 tcgagagtgatttaccctaccaggttcagctcactcccaatggagatttataaacccgtt 843  
Db 2720 TCAGAAGACGATTTACCGTACCAATTAATAAATTACTAAACCGGCGATCTCCAAACCGT 2779  
Qy 844 ggtcggttcgattttgatggacaatttagaattcccaaatgattgccaccgcaaaagtcgac 903  
Db 2780 GGAGTTACGATTTTCGACGGTCAGTTAAATTCGCAATGATAGCTCACCCGAACTGGAC 2839  
Qy 904 ccggaatccggtgaactcttcgctttaagctacgagctcgctttcaaacgcttacataaa 963  
Db 2840 CCGGTTACGAAGAGGACTTTCACGCGTTAAGCTACGACGCTGTTAAGAAACCTTACCTGAAA 2899  
Qy 964 tactcgattctcacggagcggaaactaaatcacggagctcgagattcagcttgatcag 1023  
Db 2900 TACTTCAGATTCTGCCAGACGGCGTTAAATCGCGGAATTTGGAGATCCCGCTCGAACT 2959  
Qy 1024 ccaacgagtgacagatttcgctgattacagagaacttcgctcgctacccgaccagcaaa 1083  
Db 2960 CCGACGATGATTCAGATTTTCGCTATAACGAGAGAAATTTTGTGGTGATTCCTGATCAACA 3019  
Qy 1084 gtcgttttcaagctgcgagagatgatccgcggtgggtctccggtgggtttacagacaagac 1143  
Db 3020 GTCGTGTTCAAGCTCGGCGAGATGATTTTCGCTAAATCTCCGTTGTTTTCGACGGAGAA 3079  
Qy 1144 aaggtcgcaagattcgggatttttagaataaacgcccgaagattcattcgaataaagtg 1203  
Db 3080 AAGGTTTCCCGATTGGGGATTAATGCCCAAGACCGCAGACGATTTCTCAGATAATCTGG 3139  
Qy 1204 attgatgctccagattgcttctgcttccatctctggaaacgcttgggaagagccagaaaca 1263  
Db 3140 GTGAATCTCCGGAGACGTTCTGTTTTCATCTCTGGAATGATGGGAATCGCCGAGAGCG 3199  
Qy 1264 gatgaagtcgctgtagaggttcctgtatgactcaccagactcaattttcaacagagctt 1323  
Db 3200 GAGGAGATTGTTGTTGATCGGATCGTATGTCGCGCGGCGATTCAATCTTCAACGAGAGA 3259  
Qy 1324 gacgagaatctcaagatgctcctgtctgaaatccgctcgaatctcaaaacgggtgaaaca 1383  
Db 3260 GACGAGAGCTTGAGAAGCGTTTGTTCGAGATCAGGATTAACCTCAGAACACGTAACCAACC 3319  
Qy 1384 actcgcgctccgcatcatctcaacgaagatcaacaagctcaacccctcgagcggaggtgttc 1443  
Db 3320 ACGGTCGTTCTGTTGTTGTTAAAGAGAT-----GTAATTTAGAGATTGGTATGTT 3373  
Qy 1444 aacagaaacatgctcggccgtaaaacaaatttcgcttacttggttttagccgagccgtgg 1503

Db	3374	
----	------	--

/product="putative protein"			
/protein_id="CAB78814.1"			
/db_xref="GI:7268605"			
/translation="MASGLDHFSGGIGNMLDDGSDHPIGNMLPDDDEELFSGLMDDLN LSSLPATLDDIEDYDLFGSGGLELETDYPDYSLNKGFSRMGFADSNVDNVMQNIFQNGVSIAGEHPYGEHPRTILFVRNINSNVEDSELQALFEDNPSKDVNQGTLVVFNFLAPSVNROLEINFGVYGEIKQIRETPNKRHHKFVEFFDVRSDAALAKALNRTETAGKRIKLEHSPGGARBNMLQMPLELQDDSYSLNHHVESPLASSPIGNWRNSPIDHPLOSF SKSPIGNLSPTKNIRYPESMKTASVNNDOEGRRSHLDHLFSSSSSYNNASHKASTFQQPSFGSVSSFGSLNSHPSHVETLSGSBFLWGPSSSSAWPFVNPFSNRNHRFPYSAQNSLQHLHHIGSAPGFPFRSPETSMGSVAFRGSAGNNNAQNRLETSPPNFOMLSAPRRSOLFTGNSYLPWPAATMVISIDDPLEDGSNQFDNSGNQADIKIQFOLDLSKIMRGEPTTLMIKNIPNKTNRNMLIAAIDKNSGTYYDFLYLPIDFNKNCNVCYAFINMVS PFTIALYEAFGKWKDKNSEKVASIAYRIGQKAALIAHFONSSIMNEDRCQPIVFDGSESKYPVNNLLQTHDLKMIILYFCVILNMWLTHTCCLFETRSGSEFAARASNSTG FGLFSARTHIKADSQNVNHVIASGLKRRKYQFLG"			
	exon	6377..6796	
		/gene="AT4g18120"	
		/number=1	
	intron	6797..7121	
		/gene="AT4g18120"	
		/number=1	
	exon	7122..7238	
		/gene="AT4g18120"	
		/number=2	
	intron	7239..7346	
		/gene="AT4g18120"	
		/number=2	
	exon	7347..7501	
		/gene="AT4g18120"	
		/number=3	
	intron	7502..7577	
		/gene="AT4g18120"	
		/number=3	
	exon	7578..8456	
		/gene="AT4g18120"	
		/number=4	
	intron	8457..8542	
		/gene="AT4g18120"	
		/number=4	
	exon	8543..8627	
		/gene="AT4g18120"	
		/number=5	
	intron	8628..8715	
		/gene="AT4g18120"	
		/number=5	
	exon	8716..8784	
		/gene="AT4g18120"	
		/number=6	
	intron	8785..8870	
		/gene="AT4g18120"	
		/number=6	
	exon	8871..9293	
		/gene="AT4g18120"	
		/number=7	
	exon	complement(9594..9837)	
		/gene="AT4g18130"	
	gene	complement(join(9594..9837,9944..10237,10323..11130,11238..13230))	
		/gene="AT4g18130"	
	gene	9594..13230	
		/gene="AT4g18130"	
	CDS	/gene="AT4g18130"	
		complement(join(9594..9837,9944..10237,10323..11130,11238..13230))	
		/note="Contains ATP/GTP-binding site motif A (P-loop) AA634-641; Phytochrome chromosome attachment site AA317-326 contains EST gb:T21020, AA395462"	
		/codon_start=1	
		/product="phytochrome E"	
		/protein_id="CAB78815.1"	

```
/db_xref="GI:7268606"
/translation="MGPESSSSAASNMKPPQPKNTAQYSVDAALPADPAOSIYTGKS
FNYKSVIPNHPVDEHITAYLSNIQRGLVQPGCLIAVEEPSFRLIGUSDSSDF
LGLSLPSTHSGEFKVKGILGIDARTLFTSSGASLSKAASFEISLLNPVLSHDF
TQKPFYAILHRIDAGIWMLEPAGSDPALTLAGAVSQKLAIVRAISRLQSLPGGDI
GALCDVTVDVRLTYGVORVMYOPHEDDHGEVSEIRRSLEPLVGLHYPATDIPQA
ARFLPKQNRVMI CDGNATPVKVQSEBELKPLCLVNSTLRAPHGCHTYVWANGSVA
SLALAIVYKGSXSKLGLVGHCSPRVFPFLRYACEFLMQAFLGLQMLQLASQ
LAERKAMRTQTLICMLLURDTVSAIVTOSPGIMDLVKCDGAALYKGCWLVGVTPE
SQVLDVNMVLNENHGDSTGLTSLVDAGYPGAISLDGAVCGVAAAEFSSKDYLLWF
RSNTASAIKWGAKHPKDDAGRMHPRSFYAPLEVAKRSRSLPWEISEIDAHSR
LIMRESFTSPVLSNGVARDANELTSCVEMVRVITATAPIFGVVDSSCINGWNK
KTAEMTGLVASEMCKSLADEIVOEESPALESLLCKALOGEESKVMKLKKEGQNN
HPDYSSDVCLVNSCTSDYENIIGCVFVGQDITSEKAITDRFIRLOQDRTIVQSL
NPLPIYFASDENACSEWNAAMEKLGWSKHEVIGKMLPGEVFGVCKVKQDLSLK
FLIQLYOGIADNGVPESSELYFVROEIKNPLNGIRFAHKLSESSISASOROFLETSOA
BSGSLCPLESAOSNLFTYVROEIKNPLNGIRFAHKLSESSISASOROFLETSOA
CEKQITLTIESTDLKSIIEGKLOLETFERLENILDTIISOVMIILRERNSQLAVEA
BEIKTLPLNGDRVKLQILLADLRNIVNHPPNWSVGISISPGQELSRRNGRYIHQ
FRMHPGKLSSEMLSDNFETRDGNVTPDGLGLKLSRKLLEQMGNGRVSVYREDRCFF
QVDLQVKTMLGVEGRGTGEGSSSIK"
/number=1
complement(9944..10237)
/gene="AT4g18130"
/number=2
complement(10238..10322)
/gene="AT4g18130"
/number=2
complement(10323..11130)
/gene="AT4g18130"
/number=3
complement(11131..11237)
/gene="AT4g18130"
/number=3
complement(11238..12320)
/gene="AT4g18130"
/number=4
complement(14593..14751)
/gene="AT4g18140"
/number=1
14593..16312
/gene="AT4g18140"
complement(join(14593..14751,14921..15070,15338..15639,
15915..16055,16141..16312))
/gene="AT4g18140"
complement(join(14593..14751,14921..15070,15338..15639,
15915..16055,16141..16312))
/gene="AT4g18140"
/number=1
/translation="HYA22 protein, Homo sapiens,
PIR2:JC5707
contains EST gb:R65235"
/codon_start=1
/product="putative protein"
/protein_id="CAB78816.1"
/db_xref="GI:7268607"
/translation="WKSISKDSLRKREKRVLGVCKPKHVDKVNCSFDPKISRMBELLVT
ATQCNNEEONIQGLDSTRDVNTNEPDLIVEDSSKNASNMWDTIFSPVLDDELYSET
DREPSVLLDRAEQWMLPYLEDRSASSDDVKSDEDRNINRSRKNLBAENHTAEQT
EDFPQILRNQPELADVFNYPDMQPPRSPPKRAVTFVLIWMLDILDLPDQKFSV
QFVRLDSCILSDGIYTKDLTVLGLDLAKVAIVDNCPPQYRLQINNGIPIKSWYDDPTD
DGLTLPLFLETLADANDVRPVIARFRGNKE"
/number=1
complement(14752..14920)
/gene="AT4g18140"
/number=1
complement(14921..15070)
/gene="AT4g18140"
/number=2
complement(15071..15337)
/gene="AT4g18140"
/number=2
complement(15338..15639)
```

```
/gene="AT4g18140"
/number=3
complement(15640..15914)
/gene="AT4g18140"
/number=3
complement(15915..16055)
/gene="AT4g18140"

Query Match 41.0%; Score 738.2; DB 8; Length 194143;
Best Local Similarity 68.2%; Pred. No. 3.6e-208;
Matches 1063; Conservative 0; Mismatches 478; Indels 18; Gaps 2;

QY 244 aacactaaacagatgaattgttccagagagcggcgggcagcgttggagcgcggcggag 303
DB 110145 AATCTCTCCGCTAAACATCTTCCAGAAACGGCGCGGATTCGATCGACCGCTGAG 110204

QY 304 ggtttccctgtcagccagcagaagctacaccgcgtctcctaaacggctgatcctagtgt 363
DB 110205 CGTGCAATTAATCTACAGAGCAAGATTCTCCACTTCCAAAACCGCTGATCCAGTGTT 110264

QY 364 cagatcgcggaaaatttgcctccggtgaatgaacagcgcgtccggcgttaattcttcgggtg 423
DB 110265 CAGATTGCCGGGAATTAATTCCTCCGCTACCGGAATCTTCGCTCCGCGGAAACCTCACCGTC 110324

QY 424 gtcgaaaacttcccgattccatcaaaaggagtgtatgtgcaacaggagctaacccactt 483
DB 110325 GAAGGAACAATCCCTGACTGCAATTCACGGTGTATATCCGTAAACGGCGGAATCCGATG 110384

QY 484 cagagccgggaagcaggtccaccactcttcgacgagacggtatggttcacgcccctaaa 543
DB 110385 TTTGAGCCAAACAGCTGGGACCAATTTATTCAGGAGACGGAATGTTTACCAGCAGTAA 110444

QY 544 ttcgaacacggttcagctagctacgcttgcgggtttactcagactaacccggtttgttcag 603
DB 110445 ATAACCAACGTTTCACTAGCTACGCATGCCGTTTACAAAACCCAGAGATTGTTGTTAG 110504

QY 604 gaaactcaattgggtcgaccgggtttccccaagaagcattcggtgagcttcacggccacacc 663
DB 110505 GAAAAACGATTGGGTGCGACCAAGTTTTTCCGAAAGCAATTCGGGAGCTTTCACGGTCACTCG 110564

QY 664 ggtattgccagactcatgctattctacgacagagctgagcgggtatagtcgacccggca 723
DB 110565 GGAATCGCACGTTTGATGCTGTTTACGACAGTGGGGCTTTGTGGTCTGATCAACAAACCA 110624

QY 724 caccgaaacggtgtagctaaacgcgggtttggtcctatttcaatggcgggttattggctatg 783
DB 110625 AACGCGCTCGGAGTAGCAACCGCGTTTGGTTTACTTTAATAACCGGCTTTTAGCTATG 110684

QY 784 tcggaggatgatttaccttaccaggttcagatcaactcccaatggagatttataaacccgtt 843
DB 110685 TCAGAAAGCAGATTTACCGGTACCAATTAATAAATTAATAAACCGGCGATCTCCAAACCGGTT 110744

QY 844 ggtcgggttcgatttgcggacaattagaatcccaatgattgcccaccgaaagtccac 903
DB 110745 GGACGTTACGATTTTCGACGGTCAGTTAAATCCGCAATGATAGCTACCCGCAACTGGAC 110804

QY 904 ccggaatccggtgaactttcgctttaagctacgacgctcggtttcaaaccttacctaaaa 963
DB 110805 CCGGTACGAAGGACTTCAACGGTTAGCTACGACGCTGTTAGCAAAACCTTACCTGAAA 110864

QY 964 tacttcgatttccacggacggaaactaaatcacggagcgtcgagattcagcttgatcag 1023
DB 110865 TACTTCAGATTTCTGCCAGACGGCGTTAAATCGCGGAATGGAGATCCCGCTCGAAACT 110924

QY 1024 ccaacgatgatgcagatttccgcatcacagagaacttcgctcgctaccgacagcagaa 1083
DB 110925 CCGACGATGATTCAGATTTTCGCTATTAACGGAGAAATTTGTGGTGATTCCTGTATCAACAA 110984

QY 1084 gtcgttttcaagctgcccgagatgatccgcggtgggttcctcggtgggtttacgacaagac 1143
DB 110985 GTCGTGTTCAAGCTCGCGGAGATGATTTCCGGTAAATCTCCGGTAAATCTCCGGTAAAT 111044

QY 1144 aaggtcgcaagattcgggattttagacaaatcacgccgaagatttcacgacattaaagtgg 1203
```

```

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 111045 AAGGTTTCCCGATTGGGGATAATGCCCAAGGACGGACAGAAAGCTTCTCAGATAATCTGG 111104

Qy 1204 attgatgtccagattgtttcttcttccatctctggaacgcttgggaagagccgagaaca 1263
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 111105 GTGAACCTCCCGGAGACGTTCTGTGTTTCATCTCTGGAATGCATGGGAATCGCCGAGACG 111164

Qy 1264 gatgaagtcgctgtagaggtcctgtatgactccacagactcaattttcaacgagctc 1323
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 111165 GAGGAGATTGGTGATCGGATCGTGATGTGCGCCGGCGGATTCAACTCTCAACGAGAGA 111224

Qy 1324 gacgagaatctcaagagtgctcctgtctgaaatccgcctgaaatctcaaaacccggtgaatca 1383
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 111225 GACGAGACCTTCAGAAAGCGTTTGTTCGAGATCAGGATAAACCTCAGAACACGTAATAACC 111284

Qy 1384 actcgcctccgatactctccaacgaagatcaacaagtcacacctcgaagcagggatggtc 1443
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 111285 ACGGCTCGTTCGTTGTTGTTAAACGAGGAT-----GTAAATTTAGAGATTGGTATGGTT 111338

Qy 1444 aacagaaacatgctcggcgttaaaccaaaatcgcttacttggcttttagccgagccgtgg 1503
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 111339 AACCGAAACCGTTAGGAAGAAAAACCCGGTTCGGGTTTTGGCTATTGCTTATCCTTGG 111398

Qy 1504 cctaaagtctcaggattcgctaaagtgtgatctcactactggagaagtttaagaaacatctt 1563
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 111399 CCAAAAGTTTCGGGTTTCGCTAAAGTCTGATCTTTGCACCGGTGAGATGAAAAATATATT 111458

Qy 1564 tacggcgataacccgttacggaggagagcctctgtttctcccgg-----agaa 1611
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 111459 TACGGCGGTGAGAAATATGGCGGCGAACCGTTTTCTTTCGCCGCACTCCGGTAACGGC 111518

Qy 1612 ggaggagaggagaaacgaagatacatcctctgtttcgttcacgacgagaagacatggaaa 1671
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 111519 GAAGAAAAATGAAGATGACGGGTATATATTTGTACGCTTCATGACGAAGAAACAAAGACA 111578

Qy 1672 tcggagttcacagatagtttaacgcccgttagcttagaggttgagcaacggttaaacctccg 1731
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 111579 TCAGAGCTTCAGATTATTAAACGCTGTTAAATTTAAAGCTTGAAGCTACGATTAACTACCG 111638

Qy 1732 tcaagggttcgtagcggatttcacgggtacattcatcgagccgatgatttggcgaaagca 1790
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 111639 TCTAGAGTACCGTATGGGTTTCATGGCACATTTGTGGATTGGAATGAACTCGTTGATCA 111697

```

Result No.	Query	Score	Match		Length	DB	ID	Description
			Match	Length				
1	1800	100.0	1800	22	1800	22	AAD09396	Arabidopsis thalia
2	824.6	45.8	1818	22	AAD09401	22	AAD09341	Lycopersicon escul
3	798.4	44.4	1839	22	AAD09399	22	AAD09399	Vigna unguiculata
4	738.2	41.0	1752	22	AAD09344	22	AAD09344	Arabidopsis thalia
5	606.6	33.7	1815	22	AAD09400	22	AAD09400	Zea mays neoxanthi
6	528.2	29.3	1734	22	AAD09398	22	AAD09398	Arabidopsis thalia
7	229.8	12.8	443	21	AAC56678	21	AAC56678	Eucalyptus grandis
8	204	11.3	492	22	AAC82706	22	AAC82706	Rice abscisic acid
9	188	10.4	372	21	AAC56695	21	AAC56695	Eucalyptus grandis

XX  
PA (RIKE ) RIKEN KK.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.





Db 1621 gaagacgaaggatacactctgttctcgttcacgaagagaacacatggaaatcggaggtta 1680  
 Qy 1681 cagatagttaacgcggttagcttagaggttgagcaacggttaaaacttccgtcaagggtt 1740  
 Db 1681 cagatagttaacgcggttagcttagaggttgagcaacggttaaaacttccgtcaagggtt 1740  
 Qy 1741 ccgtacggatttcacggttacattcatcgcgagccgatgatttggcgaagcaggtcgtgtga 1800  
 Db 1741 ccgtacggatttcacggttacattcatcgcgagccgatgatttggcgaagcaggtcgtgtga 1800

## RESULT 2

AAD09401

ID AAD09401 standard; cDNA; 1818 BP.

XX AC AAD09401;

XX DT 10-SEP-2001 (first entry)

XX DE Lycopersicon esculentum neoxanthin cleavage enzyme, LeNCED1 cDNA.

XX KW Tomato; neoxanthin cleavage enzyme; LeNCED1; abscisic acid; ABA;  
 KW stress tolerance; transgenic plant; plant breeding; antisense-therapy;  
 KW plant growth protectant; herbicide; ss.

XX OS Lycopersicon esculentum.

XX FH Key Location/Qualifiers  
 FT CDS 1..1818  
 FT /\*tag= a  
 FT /product= "Lycopersicon esculentum LeNCED1 protein"

XX PN EP1116794-A2.  
 XX PD 18-JUL-2001.

XX PF 11-JAN-2001; 2001EP-0300218.  
 XX PR 13-JAN-2000; 2000JP-0010056.  
 XX PR 11-JAN-2001; 2001JP-0003476.

XX PA (RIKE ) RIKEN KK.  
 XX PI Iuchi S, Kobayaashi M, Shinozaki K;  
 XX WP1; 2001-400081/43.  
 XX P-PSDB; AAE04789.

XX DR A DNA encoding a protein with a neoxanthin cleavage activity for  
 XX producing transgenic plants with improved or decreased stress tolerance  
 XX -

XX PS Claim 3; Page 67-71; 101pp; English.

XX CC The invention relates to neoxanthin cleavage enzymes and their  
 CC corresponding cDNA molecules. Neoxanthin cleavage enzyme plays a key  
 CC role in endogenous abscisic acid (ABA) biosynthesis under drought stress.  
 CC Neoxanthin cleavage enzyme is used for improving stress tolerance in a  
 CC plant when expressed in a plant cell. The invention also relates to  
 CC methods for increasing or decreasing stress tolerance in a plant by  
 CC introducing the DNA into the plant, and a transgenic plant into which a  
 CC neoxanthin cleavage enzyme is introduced. The improvement of stress  
 CC tolerance in plants is useful, for example in plant breeding. Neoxanthin  
 CC cleavage enzyme genes are useful for producing transgenic plants. An arid  
 CC land can be improved by growing transformant weed for several years and  
 CC then removing the weed by specifically lowering stress tolerance in the  
 CC weed by inducing an inducible promoter. The present cDNA sequence encodes  
 CC Lycopersicon esculentum neoxanthin cleavage enzyme, LeNCED1 protein  
 CC related to the invention.

XX SQ Sequence 1818 BP; 569 A; 383 C; 369 G; 497 T; 0 other;

Query Match 45.8%; Score 824.6; DB 22; Length 1818;  
 Best Local Similarity 69.6%; Pred. NO. 2.1e-256;  
 Matches 1134; Conservative 0; Mismatches 489; Indels 6; Gaps 1;

Qy 170 cccagctcttcattcccttaagcaatcatcaaaccttcccgcattgttgttaagcca 229  
 Db 185 cttcaaatatcaacaccaaagataatacaatttcacacccaaacaaagaaacaca 244  
 Qy 230 aagccaaagaattccaaactaaacagatgaatttgcagagagcggcggcgagcgt 289  
 Db 245 actctcttcttcaacttccaagtgaatttagtcagaagcagcagcaatggcct 304  
 Qy 290 tggacgcggcgaggggttctctgtcagccacgaagaacacacccgttctctaaacgg 349  
 Db 305 tagatgctgtagaaagtctttaactaaacatgaacttgaacaccttgcgcgaacag 364  
 Qy 350 ctgacctagtgttcagatcgccgaaatttgcgtcccggtgaatgaacagccgctccggc 409  
 Db 365 ccgacccacgaccagattctgggaatttgcgtcccggtaccggaaaaatccagtctgc 424  
 Qy 410 gtaattctccgggtgcgaaactcccgattccatcaaaaggagtgtatgtgcgaacg 469  
 Db 425 aatctcttcgggtcaccggaaaaataccaaatgttcaaggcgttacgttcgaacg 484  
 Qy 470 gagctaacccactcaacgagcgggtgacaggtcaccaacttcttcagcggagacggtatgg 529  
 Db 485 gagctaacctcttttgaaccaacgcggacacatttcttcgacggcgaggtatgg 544  
 Qy 530 ttcacgcgctcaaatcgaaacggttcagctagctacggttcgggttactcagacta 589  
 Db 545 ttcacgcgctcaaatcgaaatgggtcggttagctagcttgcggttctcatgaaacg 604  
 Qy 590 accggttcttcaggaaacgtcaattgggtcgaccggttttccccaagccatcggtagc 649  
 Db 605 agaggttcttcaagaaaaagcttgggtcgccctgtttccctaaagccattggtgaat 664  
 Qy 650 ttcacggccaaacacggtattgccgactcatgtatttaaccaggtcagatcactcccaatggag 709  
 Db 665 tacatggtcactctggaattgcaaggcttatgctgtttaaegctcgtgggctcttcggac 724  
 Qy 710 tagtcgaccggcacacggaacgggtgagtaacgcgggtttggtctatttcaatggcc 769  
 Db 725 ttgttgatcacagtaaaaggaaactggtgcaaacgcgggtttagctatttcaataacc 784  
 Qy 770 ggttattggtctatgctggaggatgatttacccttaaccaagttcagatcactcccaatggag 829  
 Db 785 gattactgctatgctgaagatgattgcttaccatgaagtaaacacccaccggcg 844  
 Qy 830 atttaaaacggttgggtcgggttcgatttggatggacaattagatcccaaatgattgcc 889  
 Db 845 atcttaaacagaggtcgatttcgatttcgacggccagctaaatccaccatgatagctc 904  
 Qy 890 accgaaagtgcaccgggaatccggtgaactcttcggttttaagctacgacgctgttcaa 949  
 Db 905 acccaagctgcaccaggttcccggtgagctatttgccttagctacgattgattcaga 964  
 Qy 950 agccttacctaaataactctccgattctcacggcggaaactaaatcacccgacgtcgaga 1009  
 Db 965 agccataactcaagctacttcgattttcaaaaaatgggaaaaatacaaatgatgtgaaa 1024  
 Qy 1010 ttcagcttgatcagccaacgatgatgcacgatttcgcgattacagagaacttcgctgcg 1069  
 Db 1025 ttccagtggaagacccaacatgatgcatttgcgaaatctactgagaacttcgctgca 1084  
 Qy 1070 tacttgacagcaactcgttttcaagctgcggagatgatcccggttgggttcgggtgg 1129  
 Db 1085 ttcctgatcaacaagctcgttttcaagatgctgaaatgatccgtggaggttcaccgggtg 1144  
 Qy 1130 tttacgacaagaacaaggtcgcaagatttcgggatttttagacaaatcacgccgaagattcat 1189  
 Db 1145 tttacgacaagaacaagtttcccgatttgggtattctctggataagtcacggaaagtgggt 1204  
 Qy 1190 cgaacattaagtggattgatgctccagattgcttctgtctccatctctggaacggttggg 1249

Db 1205 ctgatttgaaatgggtgaagtacactgattgtttctgtttccacactcgtgaatcgttggg 1264  
Qy 1250 aagaaccgaagaacagatgaagtcgtcgtagaggtcctgtatgactccacagactcaa 1309  
Db 1265 aagaagcagaacagatgaaatcgttgaattggttcgtatgacaccacagactcca 1324  
Qy 1310 ttttcaacagctgacagagaatccaagatgctcctctgaaatccgctgaatctca 1369  
Db 1325 ttttcaatgaatgatgaagggtcaatgaaggtgttttccgaaatccgtctcaatttga 1384  
Qy 1370 aaacgggtgaatcaactcgcgtcgcgcatctctcaacgaagatcaacaagtcaacctcg 1429  
Db 1395 aaacagggaatcaacaagaataatcataatcgaaaacccggatgaacaagtgaatttag 1444  
Qy 1430 aagcaggatggtcaacgaagaacatgctgcgcgttaaacacaaatctgcttacttgggtt 1489  
Db 1445 aagctggaatggtgaaccgaagaacactcgaaggaacacagatgattcttattggcta 1504  
Qy 1490 tagcgaagcctggcctaaagtctcaggattcgcgtctaaagtgtatctcactactggagaag 1549  
Db 1505 tgcctgaaccatggccaaagtcttctgttttgcaaaagtataacctgttcacoggtgaag 1564  
Qy 1550 ttaagaacatctttacggcgataaccgttacggaggagagcctctgtttctccocggag 1609  
Db 1565 ttgagaattcatttatggtgacaacaataatggtgggaacacctttttttaccagaag 1624  
Qy 1610 -----aaggaggagagaagaacagatatacattctctgtttcgttcacacagagaaga 1663  
Db 1625 accccaacagcaaggagaagacgatggttatatttttgcgtttcgttcacgatgagaag 1684  
Qy 1664 catggaatccgagttacagatagtttaacgcgttagcttagaggttgaagcaacggtta 1723  
Db 1685 aatggaatcgaactgcaaatgtttaacgcaatgagtttgaagtggaggcaactgtga 1744  
Qy 1724 aacttccgtcaaggttccgtacgatttccaggttacatttcacgtgagccgatgatttgg 1783  
Db 1745 agcttccatcaagaagtctcttatgatttcatggaacattcataaacgccaatgatttgg 1804  
Qy 1784 cgaagcagg 1792  
Db 1805 caaatcagg 1813

RESULT 3  
AAD09399 standard; cDNA; 1839 BP.

XX AC AAD09399;

XX DT 10-SEP-2001 (first entry)

XX DE Vigna unguiculata neoxanthin cleavage enzyme, CPRD65 cDNA.

XX KW Cowpea; neoxanthin cleavage enzyme; abscisic acid; ABA; herbicide;  
XX KW stress tolerance; transgenic plant; plant breeding; antisense-therapy;  
XX KW plant growth protectant; Cowpea Responsive to Dehydration; CPRD65; ss.

XX OS Vigna unguiculata.

XX FH Key Location/Qualifiers

XX FT CDS 1..1839

XX FT /\*tag= a

XX FT /product= "Vigna unguiculata CPRD65 protein"

XX PN EP1116794-A2.

XX PD 18-JUL-2001.

XX PF 11-JAN-2001; 2001EP-0300218.

XX PR 13-JAN-2000; 2000JP-0010056.

XX PR 11-JAN-2001; 2001JP-0003476.

XX (RIKE ) RIKEN KK.  
PA  
XX Iuchi S, Kobayashi M, Shinozaki K;  
XX WPI; 2001-400081/43.  
DR P-PSDB; AAE04787.  
XX

PT A DNA encoding a protein with a neoxanthin cleavage activity for  
PT producing transgenic plants with improved or decreased stress tolerance  
PT  
XX Claim 3; Page 53-56; 101pp; English.  
XX

XX The invention relates to neoxanthin cleavage enzymes and their  
CC corresponding cDNA molecules. Neoxanthin cleavage enzyme plays a key  
CC role in endogenous abscisic acid (ABA) biosynthesis under drought stress.  
CC Neoxanthin cleavage enzyme is used for improving stress tolerance in a  
CC plant when expressed in a plant cell. The invention also relates to  
CC methods for increasing or decreasing stress tolerance in a plant by  
CC introducing the DNA into the plant, and a transgenic plant into which a  
CC neoxanthin cleavage enzyme is introduced. The improvement of stress  
CC tolerance in plants is useful, for example in plant breeding. Neoxanthin  
CC cleavage enzyme genes are useful for producing transgenic plants. An arid  
CC land can be improved by growing transformant weed for several years and  
CC then removing the weed by specifically lowering stress tolerance in the  
CC weed by inducing an inducible promoter. The present cDNA sequence encodes  
CC Vigna unguiculata neoxanthin cleavage enzyme, CPRD65 (Cowpea Responsive  
CC to Dehydration) protein. CPRD65 gene is isolated from cowpea plant.  
XX  
SQ Sequence 1839 BP; 447 A; 595 C; 442 G; 355 T; 0 other;

Query Match 44.4%; Score 798.4; DB 22; Length 1839;  
Best Local Similarity 70.4%; Pred. No. 6.7e-248;  
Matches 1098; Conservative 0; Mismatches 456; Indels 6; Gaps 2;

Qy 233 ccaagaatccaactaaacagatgaattgttccagagacgcgcgccagcgttgg 292  
Db 281 ccaaccaccattacctcaaaatggaactttctccgaaagccgtgcacgccttgg 340  
Qy 293 acgcgccggagggtttcttctgcagccacgagaaagctacacccgtcttctaaaacgctg 352  
Db 341 acctggtcgaacggcgtcgtctcgcaagcgcaaacacccgctcccaaaacggcg 400  
Qy 353 atcctagtgtcagatcgccggaattttgtccggtgaaatgaacagccctccgggcta 412  
Db 401 acccaggggtccaaatcgccgggaacttcgcccgtgcccagcatgctcgatcaag 460  
Qy 413 atctccggtggtcggaacttcccgattccatccaaaggagtgtgtgcgaacggag 472  
Db 461 gactcccggtggtcggaacttcccaaatcccaatgcatgacggtgacgtgcgaacgg 520  
Qy 473 ctaccacttcacgagccggtgacaggtcaccacttcttcgacggagagcgttattggttc 532  
Db 521 ccaatcgctctacagcctgtgcccgggaccacttcttcgacggcagcgtatggtcc 580  
Qy 533 acgcccgaattcgaacacgggttcagctagctacgttgcggtttactcagactaac 592  
Db 581 acgctgtgaagtccagaaacggccgcagctacgctgcgcttccacgagacgcagc 640  
Qy 593 ggtttgttcaggaacgtcaattgggtgcacgggtttcccaaaagccatcgttgcgttc 652  
Db 641 gtctctcgaggagaataatcttagccgcgcggtgttcccgaaagccatcgttgcgttc 700  
Qy 653 acggccacacccggtattgcccgactcatgctattctacgcagagctgcagccggtatag 712  
Db 701 acggccactcggcatcgcggtctctcttctacgcgcggtctctctcgggtctcg 760  
Qy 713 tcgacccggcacacgggaacccggtttagctaaacgcggtttgggtctatttaattggcgg 772  
Db 761 ttgatgggtccacgggcatggcggtggcgaacgcggtctcgtctacttcaacaaccacc 820

QY 773 tattgctatgtcggagatgattttacattaccagttcagatcactccaatgagatt 832  
Db 821 tcttgccatgtccgaagcagattttaccctaccacgtgagaatcacccctaacgcgcact 880  
QY 833 taaaaaccgttggtcgttcctgatttgatggacaattagaatccacaatgattgccacc 892  
Db 881 taaccaccgttggtcgttcctgatttgatggacaattagaatccacaatgattgccacc 940  
QY 893 cgaagtcgaccgggaatccggtgaaactttcgctttaagctacgagctcggtttcaaacg 952  
Db 941 cgaactggaccgccgtcgagcgccactccacgctcagctacgactcattcagaagc 1000  
QY 953 cttaactaaatcttcctgatttcacggagcgaactaaatcacggcgcgtcgagattc 1012  
Db 1001 cttaactcaagtacttccttcctcccgagcggtcgaagtcctcccgagcggtggaatcc 1060  
QY 1013 agcttgatcagcaacgagatgacgagtttcgctgattacagagaaactcgtcgtgac 1072  
Db 1061 cctgaaggagccaccatgacgagtttcgctgatttcgctgatttcgctgctgctcc 1120  
QY 1073 ctgaccagcaagtctgtttcaagctgcggagatgacgcggtggttcctcggtggttt 1132  
Db 1121 cggaccagcaggtggtctcaactaacggagatgacacggcggttcccggtggtct 1180  
QY 1133 acgaacgaacaggtgcgaagttcggagttttgagacaataacgccgaagattcatoga 1192  
Db 1181 acgaacgaacaaactcaggttggatttctgcagaagaatgcgaaggcgcgaatg 1240  
QY 1193 acattaaaggattgattcctcagattctctgcttcacatctcggaaacgttgggaag 1252  
Db 1241 cgatcggtggtacgcgcgcggattgttctgcttcacatctcggaaacgttgggaag 1300  
QY 1253 agccgaacacagatgaagtcgtgattgaggtcgtgattgacacccagactcaattt 1312  
Db 1301 agccgaacacagaggttgggtgattgaggtcgtgattgacacccagactcaattt 1360  
QY 1313 tcaacgagttcagagaatctcaagagttcgtgctgtaaatccgctgaattcctcaaaa 1372  
Db 1361 tcaacgagttcagagaatctcaagagttcgtgctgtaaatccgctgaattcctcaaaa 1420  
QY 1373 cgggtgaactcaactcgcctcgcattctccaaagaaatcaaacgtcaacctcgaag 1432  
Db 1421 cgggaagtcctcctcgcctcgcattctccaaagaaatcaaacgtcaacctcgaag 1477  
QY 1433 cagggtggtcgaacgaacatgctcgcgcgttaaaccaaatcgttacttgggttttag 1492  
Db 1478 cgggtggtggaacgaacatgctcgcgcgttaaaccaaatcgttacttgggttttag 1537  
QY 1493 cgaagcgttggtcctaaagtctcagattcgtcgaagttgattcactactcagagaagtt 1552  
Db 1538 cgaagcgttggtcctaaagtctcagattcgtcgaagttgattcactactcagagaagtt 1597  
QY 1553 aaaaacatcttcacggcgaataacgttacggagagagcctcgttctcccggaag 1612  
Db 1598 aagaatcagatgaggaagaagaggttcggtggagagcctcgttctcccggaag 1654  
QY 1613 gagggagagaacgaagatacctcctcgttgcgttcacgacgagagacatggaat 1672  
Db 1655 gccaaaagaacagatgggtatattcggcattcgtgcacgacgagaagaatggaat 1714  
QY 1673 cgaagttacagatagtaacgcgttagcttagaggttgaaacacggttaaacctcgt 1732  
Db 1715 cgaagttacagatagtaacgcgttagcttagaggttgaaacacggttaaacctcgt 1774  
QY 1733 caaggttcggtacgatttcaggtacattcattcggagcagatatttggcgaagcgg 1792  
Db 1775 ctctgttccctacaggtttcttggaactttcattcattccaaagatttgaggaaacag 1834

RESULT 4

AAD09394

ID AAD09394 standard; cDNA; 1752 BP.

XX

AC AAD09394;

XX 10-SEP-2001 (first entry)

XX Arabidopsis thaliana neoxanthin cleavage enzyme, AtNCED1 cDNA.

XX Neoxanthin cleavage enzyme; AtNCED1; abscisic acid; ABA; herbicide;  
stress tolerance; transgenic plant; plant breeding; antisense-therapy;  
plant growth protectant; ss.

XX Arabidopsis thaliana.

FH Key Location/Qualifiers  
CDS 1..1752  
FT /\*tag= a

FT /product= "Arabidopsis thaliana AtNCED1 protein"

XX EP1116794-A2.

XX 18-JUL-2001.

XX 11-JAN-2001; 2001EP-0300218.

XX 13-JAN-2000; 2000JP-0010056.

PR 11-JAN-2001; 2001JP-0003476.

XX (RIKE ) RIKEN KK.

XX Iuchi S, Kobayashi M, Shinozaki K;

XX WPI; 2001-400081/43.

DR P-PSDB; AAE04782.

XX A DNA encoding a protein with a neoxanthin cleavage activity for  
producing transgenic plants with improved or decreased stress tolerance

Claim 3; Page 18-22; 101pp; English.

The invention relates to neoxanthin cleavage enzymes and their  
corresponding cDNA molecules. Neoxanthin cleavage enzyme plays a key  
role in endogenous abscisic acid (ABA) biosynthesis under drought stress.  
Neoxanthin cleavage enzyme is used for improving stress tolerance in a  
plant when expressed in a plant cell. The invention also relates to  
methods for increasing or decreasing stress tolerance in a plant by  
introducing the DNA into the plant, and a transgenic plant into which a  
neoxanthin cleavage enzyme is introduced. The improvement of stress  
tolerance in plants is useful, for example in plant breeding. Neoxanthin  
cleavage enzyme genes are useful for producing transgenic plants. An arid  
land can be improved by growing transgenic plants for several years and  
then removing the weed by specifically lowering stress tolerance in the  
weed by inducing an inducible promoter. The present cDNA sequence encodes  
Arabidopsis thaliana neoxanthin cleavage enzyme, AtNCED1 protein.  
The AtNCED1 cDNA is obtained from an Arabidopsis plant-derived cDNA  
library using a cDNA of the cPRD65 (Cowpea Responsive to Dehydration)  
gene isolated from cowpea plant as a probe.

SQ Sequence 1752 BP; 484 A; 396 C; 428 G; 444 T; 0 other;

Query Match 41.0%; Score 738.2; DB 22; Length 1752;  
Best Local Similarity 68.2%; Pred. No. 2.2e-228;  
Matches 1063; Conservative 0; Mismatches 478; Indels 18; Gaps 2;

QY 244 aacactaaacagatgaattttccagagagcggcgccgagcgttggacgcggcggag 303

Db 193 aatcctctccggcgaacatcttcagaaagcggcgccgattgcgacgcggcgtgag 252

QY 304 ggtttccttgcagcagcagaagctacaccgcttctctaaaacggctgacctagttt 363

Db 253 cgtgcattaatctcacagcagcagaagattctcacttcccaaacgcgtgacctggtt 312

QY 364 cagatcgccggaattttgctccgggtgaatgaacagccgcgtccggtatcttccggtg 423

Db 313 cagattcgcgggaattattcccggtaccggaattcttcgctcgcggcgaacctcaacgtc 372  
Qy 424 gtcggaataactcccgattccatcaaaagagtgtatgctgcgaacggagctaaacccactt 483  
Db 373 gaaggaaacaatccctgactgactgacgggtttatatccgtaacggcgcgaatccgatg 432  
Qy 484 cacgagccgggtgacaggttcacacactcttcgacgagagcggatggttcacgcgctcaaa 543  
Db 433 tttagcacaacagctcggcaccatttatcgcgagagcgaatgggttcacgcagttaaa 492  
Qy 544 ttcgaacacgggttcagctcagctacgcttcgcgggtttactcagactaacgggtttggttcag 603  
Db 493 ataaccaacgggttcagctcagctacgcatcgccgggtttacaaaacccgagagattgggttcag 552  
Qy 604 gaacgtcaattgggttcgacccgggtttcccaaaagccatcggtgagcttcacggccacacc 663  
Db 553 gaataacgattgggttcgaccagttttccgaaagcaatcggcgagcttcacggctcaactcg 612  
Qy 664 ggtattgccgactcaatctattcaccgagagctgcagccggtatagtcgaccggcga 723  
Db 613 ggaatcgcacgttgatgctgtttacgcacgtgggctttgtggtctgatcaacaaccaa 672  
Qy 724 caccgaaacgggtgtagctaaacgcgggtttggtctatttcaatggccgggttattgggtatg 783  
Db 673 aacggcgtcggagtagcaaacgcgggtttggtttactttaataacgggttttagctatg 732  
Qy 784 tcggaggatgatttaccctaccaggttcagatcactcccaatggagatttaaaaacggtt 843  
Db 733 tcagaagacgatttaccgtaccatataaaattactcaaacggcgatctccaaacggt 792  
Qy 844 ggtcgggttcgatttgatggacaattagaatccacaatgattgccaccgaaagtcgac 903  
Db 793 ggaacttaccgatttcgcggtcagtttaaaatccgcaatgtagctcacccgaaactggac 852  
Qy 904 ccggaatccggtggaactcttcggttaagctacgacgtcggtttcaaaagccttacctaaaa 963  
Db 853 ccggttacgaaggacttcacgcgttaagctacgacgtcggttaagaacaccttacctgaaa 912  
Qy 964 tacttccgattctcaccggacggaaactaaatcaccggacgtcgagattcagcttgatcag 1023  
Db 913 tacttccgattctcaccgacggcgcttaaatccgcgggaattggagatcccgctcgaaact 972  
Qy 1024 ccaacgattgacagatttccggtattcagagaacttcgctcgtcgtcactgacacagaa 1083  
Db 973 ccgacgattgattcagatttcgctataacggaagaatttgggtgattcccgatcaacaa 1032  
Qy 1084 gtcggtttcaagctccgagatgatccggtgggttcggttcggttttacgacaagaac 1143  
Db 1033 gtcggttcaagctccgagatgattccggttaaatcccggttgggttttcgacggagaa 1092  
Qy 1144 aaggtcgaagattcgggatttttagacaataacgccgaagatttcacgaacattaaagtgg 1203  
Db 1093 aaggtttccgattggggataatgccgaagacgcgacagaagcttctcagataatctgg 1152  
Qy 1204 attgattccagattgcttctgcttcattctcgaacggttgggaagagccgagaaac 1263  
Db 1153 gtgaactctcggagacattctgtttctcattctcgtggaatgcatgggaatcccgagacg 1212  
Qy 1264 gatgaagtcgctgctgatagggtccctgtagctccaccagactcaattttcaacgagctt 1323  
Db 1213 gaggagattgggtgatcggatcgtgtagtgcgcggcggttcaatcttcaacgagaga 1272  
Qy 1324 gacgagaatctcaagagtgtcctgtctgaaatccgcctgaatctcaaaacccggtgaatca 1383  
Db 1273 gacgagcttgagaagcgttttgcggagatcaggataaacctcagaacacgtaaaaacc 1332  
Qy 1384 actccgctcggatcaatctcaacgaagaatcaaacggtcaacctcgaagcggagatgctc 1443  
Db 1333 acgctcgttctggttgggttaacgagat-----gtaatattagagattgggtatggt 1386  
Qy 1444 aacagaaacatgctcggccgttaaaaccaaattcgttacttgggtttagcggagcgtgg 1503

Db 1387 aaccggaaacgggttaggaagaagaaaaaacccggttcgctgttttggctattgcttaccctgg 1446  
Qy 1504 cctaaaagttccagattcgtctaaagttgatctcaactactggagaagttaagaaaaacatctt 1563  
Db 1447 ccaaaagttccgggttcgctaaaggtcgatctttgcaccgggtgagatgaaaaaatatt 1506  
Qy 1564 tacggcgataaacggttacgagagagacccctctgtttctccccgg-----agaa 1611  
Db 1507 tacggcggtgagaaatattgcccgaacccgttttcttccccggaactccggttaacggc 1566  
Qy 1612 ggagagaggaagcgaagatacatctctgtttcttccacgacgagaagacatggaaa 1671  
Db 1567 gaagaaaatgaagatgacgggttatatttgcacgttcatgacgagaacaaagaca 1626  
Qy 1672 tcggagttacagatagtttaacgcggttagcttagaggttgaagcaacgggttaaaactccg 1731  
Db 1627 tcagagttcagattattaacgctgttaatttaagttgaagctacgattaaactaccg 1686  
Qy 1732 tcaaggggtccgtacgatttcacggttacatttcacggacgagcgtgatttggcgaaaga 1790  
Db 1687 tctagatccggtatgggtttcattcggcacatttgggattcgaatgaactcgttgatca 1745

RESULT 5  
AAD09400  
ID AAD09400 standard; cDNA; 1815 BP.  
XX  
AC AAD09400;  
XX  
DT 10-SEP-2001 (first entry)  
XX  
DE Zea mays neoxanthin cleavage enzyme, VP14 cDNA.  
XX  
KW Maize; neoxanthin cleavage enzyme; VP14; abscisic acid; ABA; herbicide;  
KW stress tolerance; transgenic plant; plant breeding; antisense-therapy;  
KW plant growth protectant; ss.  
XX  
OS Zea mays.  
XX  
FH Key Location/Qualifiers  
CDS 1..1815  
FT /\*tag= a  
FT /product= "Zea mays VP14 protein"  
XX  
EP EPI116794-A2.  
XX  
PD 18-JUL-2001.  
XX  
PF 11-JAN-2001; 2001EP-0300218.  
XX  
PR 13-JAN-2000; 2000JP-0010056.  
PR 11-JAN-2001; 2001JP-0003476.  
XX  
PA (RIKE ) RIKEN KK.  
XX  
PI Iuchi S, Kobayashi M, Shinozaki K;  
XX  
DR WPI; 2001-400081/43.  
DR P-PSDB; AAE04788.  
XX  
PT A DNA encoding a protein with a neoxanthin cleavage activity for  
PT producing transgenic plants with improved or decreased stress tolerance  
XX  
XX Claim 3; Page 60-64; 101pp; English.  
XX  
CC The invention relates to neoxanthin cleavage enzymes and their  
CC corresponding cDNA molecules. Neoxanthin cleavage enzyme plays a key  
CC role in endogenous abscisic acid (ABA) biosynthesis under drought stress.  
CC Neoxanthin cleavage enzyme is used for improving stress tolerance in a  
CC plant when expressed in a plant cell. The invention also relates to  
CC methods for increasing or decreasing stress tolerance in a plant by  
CC introducing the DNA into the plant, and a transgenic plant into which a







Qy 1794 cgtgtga 1800  
 Db 1728 ttctcga 1734

## RESULT 7

AAC56678  
 ID AAC56678 standard; DNA: 443 BP.

XX  
 AC AAC56678;

XX  
 DT 25-JAN-2001 (first entry)

XX  
 DE Eucalyptus grandis transcription factor DNA sequence #549.

XX Plant; transcription factor; gene expression; eucalyptus; pine; acacia;  
 KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;  
 KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;  
 KW homeodomain zipper; LIM domain; AP2; EREBs; zinc finger domain;  
 KW type 2 Cys2His2; CCAAT box element; MYB; ss.

XX Eucalyptus grandis.

XX WO200053724-A2.

XX PD 14-SEP-2000.

XX PF 09-MAR-2000; 2000WO-US06112.

XX PR 11-MAR-1999; 99US-0266513.

XX PR 18-AUG-1999; 99US-0149485.

XX PA (GENE-) GENESIS RES & DEV CORP LTD.

XX PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.

XX PI Wood M, McGrath A, Shenk MA, Glenn M;

XX DR WPI; 2000-579369/54.

XX New isolated polynucleotide encoding a plant transcription factor for  
 PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,  
 PT having modified gene expression or modified activity of a polypeptide

XX PS Claim 1; Page 490; 747pp; English.

XX The present invention relates to novel plant transcription factors from  
 CC Eucalyptus grandis or Pinus radiata. The present sequence is the coding  
 CC sequence for one such transcription factor. The transcription factor may  
 CC be used to produce a plant having modified gene expression such as a  
 CC woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or  
 CC mahogany species or to modify the activity of a polypeptide in a plant.  
 CC The transcription factors of the present invention are members from the  
 CC following families of regulatory proteins: bZIP, bZIP family of G-box  
 CC binding factors, basic helix-loop-helix zipper,  
 CC homeotic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2  
 CC and EREBs, zinc finger domains of type 2 Cys2His2, CCAAT box elements  
 CC and MYB.

XX Sequence 443 BP; 71 A; 177 C; 117 G; 77 T; 1 other;

Query Match 12.8%; Score 229.8; DB 21; Length 443;  
 Best Local Similarity 69.9%; Pred No. 8.5e-64;  
 Matches 309; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

Qy 620 gaccggtttcccaagccatcggtgagcttcacggccacacccggtatgccgactca 679

Db 1 gcccgcttcncaaggccatcggtgagcttcacggccacacccggtatgccgactca 60

Qy 680 tgcatttaccgacgagctgcgacgggtatagtcgaccccgacacggaaacgggtgag 739

Db 61 tgcatttaccgacgagcttcgacctgctgcgaccccggaatggcatggcgctcg 120  
 Qy 740 ctaacgcccgtttggtctatttcaatggccggttattggctatgctggaggatgatttac 799  
 Db 121 cgaacgcccgttcgtgtaactcgacggccacctctcgcatgctcggaggacgacctcc 180  
 Qy 800 cttaccaggttcagatcactcctccaatggagattttaaaaacggttgggttcgattttg 859  
 Db 181 cctaccaggtgcgctcagcgctccggcagctcgagacctgcggcgtcagcacttcg 240  
 Qy 860 atggacaattagattcccaatgattgcccccgaagtcgacccggaatccggtgaac 919  
 Db 241 ccggccagctgactctccgcatgctgcaccccggaagtcgacccggttcggtcgaga 300  
 Qy 920 tcttcgcttaagctacgacgtcgtttcaagccttacctaaatacttccgattctcac 979  
 Db 301 tgttcgctcagctacgacgtcgtccggaagcgtacctcaagtacttccgattctcca 360  
 Qy 980 cgaacggaactaaatcacggacgtcgagattcagcttgatcagcgaacgatgacgacg 1039  
 Db 361 aggaacgagagaagtcccccgacgtcgagatccccctggtgagccgacctgatgacg 420  
 Qy 1040 atttcgcatcagagaactt 1061  
 Db 421 atttcgcatcagagaactt 442

## RESULT 8

AAC82706/C  
 ID AAC82706 standard; DNA: 492 BP.

XX AC AAC82706;

XX DT 15-MAR-2001 (first entry)

XX DE Rice abscisic acid synthesis associated DNA SEQ ID NO 5.

XX KW Rice; abscisic acid synthesis; VP14-like gene; drought resistance; ds.

XX OS Oryza sativa.

XX PN WO200071727-A1.

XX PD 30-NOV-2000.

XX PF 25-MAY-1999; 99WO-JP02734.

XX PR 25-MAY-1999; 99WO-JP02734.

XX PA (NORQ) JAPAN MIN AGRIC FORESTRY & FISHERIES.

XX PA (BIOO-) BIO-ORIENTED TECHNOLOGY RES ADVANCEMENT.

XX PI Hirochika H, Sakamoto K;

XX DR WPI; 2001-032042/04.

XX Oligonucleotide encoding gene for regulating abscisic acid synthesis in  
 PT plants, useful for constructing e.g. genetically-modified rice with  
 PT drought resistance and ear-germination resistance

XX PS Disclosure; Page 50; 55pp; Japanese.

XX This invention describes a novel polynucleotide sequence (I) which  
 CC encodes a protein capable of regulating the synthesis of abscisic acid.  
 CC The invention also describes (1) an oligonucleotide encoding a protein  
 CC of a gene obtained by controlling the expression of a VP14-like gene;  
 CC (2) a vector containing the oligonucleotide ligated operably to the  
 CC regulation sequence; (3) a plant transformed with the vector; and  
 CC (4) a method for regulating abscisic acid synthesis in a plant including  
 CC the transfer of the above oligonucleotide to it. The gene is useful for  
 CC constructing drought resistant rice.

XX SQ Sequence 492 BP; 90 A; 126 C; 186 G; 88 T; 2 other;

```
Query Match      11.3%; Score 204; DB 22; Length 492;
Best Local Similarity 69.7%; Pred. No. 2.1e-55;
Matches 304; Conservative 0; Mismatches 130; Indels 2; Gaps 2;

QY 627 tttcccaaacccatcggtgagcttcacggccacacc-ggtattgcccgaactatgtat 685
DB 435 TTTTCTAAGCGTAGGTGAGCTCCATGGCCACTCCGGGATCGCGCCTTGCTCTGT 376
QY 686 tctacgccagagctgacgggtatagtcgaccgggcacacgggaaccgggtagctaacg 745
DB 375 TCTAGCGCGCGCGC-CTCGGCTCTCTGACCCGCTCACACGGCACCGGCGCTCGCCAAAG 317
QY 746 ccggtttgggtatttcaatggccggttatggcttatgctcgagagatgattaccttacc 805
DB 316 CCGGCTCTATCTACTTCAACGGCAGGCTCTCGCCATGTCGGAGGACGACCTCCCTTACC 257
QY 806 aagttcagatcactcccaatgagatttataaacggttggctggttcgattttgatggac 865
DB 256 AGGTGCGGCTACCGCCGACGGCGACCTCGAGACGTCGGCGGCTACGACTTCGACGGGC 197
QY 866 aattagaatccacaatgattgccaccgcgaagtcgacccgggaatccggtgaactcttcg 925
DB 196 AGCTCGGCTGCGCATGATCGGCACACCCCAAGCTCGACCCGCGCACCGGAGCTCCACG 137
QY 926 ctttaagtacgagctgcttcaagcctttacctaataatttcgatttcacccggagc 985
DB 136 CGCTCAGCTACGACGTGATCAAGAACGCGTACTCTCAAGTACTTCTACTTCGCGCCGACG 77
QY 986 gaactaaatccggacgctcgagattcgcttgatcgacccaacgatgatgcagatttcg 1045
DB 76 GCACCAAGTCGGCGACGTCGAGATCCCGCTCGACACCCACCATGATGATCCAGACTTCC 17
QY 1046 cgattacagagaactt 1061
DB 16 GAATTACTGAGTTCTT 1

RESULT 9
AAC56695
ID AAC56695 standard; DNA; 372 BP.
XX
AC AAC56695;
XX
DT 25-JAN-2001 (first entry)
XX
DE Eucalyptus grandis transcription factor DNA sequence #566.
XX
KW Plant; transcription factor; gene expression; eucalyptus; pine; acacia;
KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;
KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;
KW homeodomain zipper; LIM domain; AP2; EREBs; zinc finger domain;
KW type 2 Cys2His2; CCAAT box element; MYB; ss.
XX
OS Eucalyptus grandis.
XX
FN WO200053724-A2.
XX
PD 14-SEP-2000.
XX
PF 09-MAR-2000; 2000WO-US06112.
XX
PR 11-MAR-1999; 99US-0266513.
PR 18-AUG-1999; 99US-0149485.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI Wood M, McGrath A, Shenk MA, Glenn M;
XX WPI; 2000-579369/54.
XX
```

New isolated polynucleotide encoding a plant transcription factor for producing a plant e.g. a woody plant, preferably eucalyptus or pine, having modified gene expression or modified activity of a polypeptide

Claim 1; Page 494; 747pp; English.

The present invention relates to novel plant transcription factors from *Eucalyptus grandis* or *Pinus radiata*. The present sequence is the coding sequence for one such transcription factor. The transcription factor may be used to produce a plant having modified gene expression such as a woody plant e.g. a *eucalyptus*, pine, acacia, poplar, sweetgum, teak, or mahogany species or to modify the activity of a polypeptide in a plant. The transcription factors of the present invention are members from the following families of regulatory proteins: bZIP, bZIP family of G-box binding factors, basic helix-loop-helix zipper, homeotic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2 and EREBs, zinc finger domains of type 2 Cys2His2, CCAAT box elements and MYB.

Sequence 372 BP; 60 A; 148 C; 101 G; 63 T; 0 other;

Query Match 10.4%; Score 188; DB 21; Length 372;  
Best Local Similarity 69.1%; Pred. No. 2.8e-50;  
Matches 257; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 639 catcggtgagcttcacggccacacccggtattgcccgaactatgctatttcacgacgagc 698  
DB 1 catcgcgagctccacggccacactccggcatcgcggtcatgctcttctacgcccag 60  
QY 699 tgcagccggtatagtcgacccgggcacacgggaacgggtgtagctaacccgggttggctta 758  
DB 61 cctcttcggcctcgctgacaccgggaatggatggcggtcggaacccggcctcgtga 120  
QY 759 ttccaatggcggttattggtatgctcgagagatgattaccttaccgaagtctagatcac 818  
DB 121 ctccgagggcaccctctcgatgctcgaggagacactccctaccacacgtgcgctcac 180  
QY 819 tcccaatggagagatttaaaacccgttgctcggttcgattttgatggacaattagatccac 878  
DB 181 gcgtccggcgacctcgagacccgtcgccgctacgacttcgcccggcagctcgactctcc 240  
QY 879 aatgattgccccccgaaagtcgacccgggaatccgggtgaactcttcgctttaaagctcga 938  
DB 241 gatgatcgccccaccggaagatcgaccgggttcctcgcgagatgctccctcagctcaga 300  
QY 939 cgctgcttcaaaagccttaccataaatacttccgattctcccggaacggaactaaatcacc 998  
DB 301 cgctcgccggaagcgtacctcaagtacttccgatttcccaaggacggcgagaagtcccc 360  
QY 999 ggacgtcgagat 1010  
DB 361 cgacgtcgagat 372

RESULT 10  
AAC56548  
ID AAC56548 standard; DNA; 325 BP.  
XX  
AC AAC56548;  
XX  
DT 25-JAN-2001 (first entry)  
XX  
DE Eucalyptus grandis transcription factor DNA sequence #419.  
XX  
KW Plant; transcription factor; gene expression; eucalyptus; pine; acacia;  
KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;  
KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;  
KW homeodomain zipper; LIM domain; AP2; EREBs; zinc finger domain;  
KW type 2 Cys2His2; CCAAT box element; MYB; ss.  
XX  
OS Eucalyptus grandis.

XX WO200053724-A2.  
 XX PD 14-SEP-2000.  
 XX PF 09-MAR-2000; 2000WO-US06112.  
 XX PR 11-MAR-1999; 99US-0266513.  
 XX PR 18-AUG-1999; 99US-0149485.  
 XX (GENE-) GENESIS RES & DEV CORP LTD.  
 PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.  
 XX PI Wood M, McGrath A, Shenk MA, Glenn M;  
 XX DR WPI; 2000-579369/54.  
 XX New isolated polynucleotide encoding a plant transcription factor for  
 PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,  
 PT having modified gene expression or modified activity of a polypeptide  
 PT  
 XX  
 XX  
 XX Claim 1; Page 461; 747pp; English.  
 CC The present invention relates to novel plant transcription factors from  
 CC Eucalyptus grandis or Pinus radiata. The present sequence is the coding  
 CC sequence for one such transcription factor. The transcription factor may  
 CC be used to produce a plant having modified gene expression such as a  
 CC woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or  
 CC mahogany species or to modify the activity of a polypeptide in a plant.  
 CC The transcription factors of the present invention are members from the  
 CC following families of regulatory proteins: bZIP, bZIP family of G-box  
 CC binding factors, basic helix-loop-helix zipper,  
 CC homeotic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2  
 CC and EREBs, zinc finger domains of type 2 Cys2His2, CCAAT box elements  
 CC and MYB.  
 XX  
 XX Sequence 325 BP; 45 A; 128 C; 100 G; 52 T; 0 other;

Query Match 9.5%; Score 171.8; DB 21; Length 325;  
 Best Local Similarity 71.2%; Pred. No. 4.6e-45;  
 Matches 227; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

Qy 451 ggagtgatgtgcgaacggagcgtacacacgttcacgagccggtgacaggtcaccacttc 510  
 Db 3 ggggtgactgcgcgaacggcgcacccgcgtccacgagccggtgcgggacacacttg 62  
 Qy 511 ttcgacgagacggatggttcacgcccgtcaaatcgaacacggttcagctacgct 570  
 Db 63 ttcgacggcgaacgatgatccacgcgtccggttctccggcggtcagtgagctacgcc 122  
 Qy 571 tgcgggtttactcagactaacgggtttgttcaggaaactcaattgggtcgaacgggttttc 630  
 Db 123 tgcgggttcacgagacgaacacgctgatccaggaaacggggtccgcccgcctcttc 182  
 Qy 631 cccaaagcctcgttgagcttcacggccacacacggtattgcccgactatgctattctac 690  
 Db 183 cccaaggcctcggcagcttcacggccacacacggtcggcgcgtcagctctctctac 242  
 Qy 691 gccagagctgcagccggttatagtcacccgggcacacggaacgggtgtagctaacgccggt 750  
 Db 243 gcccgccgctcttcggctcgtcgcacacacgtaatggcgtggcggtcgcgaacgcgggc 302  
 Qy 751 ttggtctatttcaatggcc 769  
 Db 303 ctctgtacttcgacggcc 321

RESULT 11  
 AAC42989  
 ID AAC42989 standard; DNA; 1788 BP.  
 XX

AC AAC42989;  
 XX 17-OCT-2000 (first entry)  
 DT XX Arabidopsis thaliana DNA fragment SEQ ID NO: 37611.  
 DE XX Hybridisation assay; genetic mapping; gene expression control;  
 XX protein identification; signal transduction pathway;  
 KW metabolic pathway; promoter; termination sequence; ss.  
 KW Arabidopsis thaliana.  
 OS Arabidopsis thaliana.  
 XX EP1033405-A2.  
 XX PD 06-SEP-2000.  
 XX PF 25-FEB-2000; 2000EP-0301439.  
 XX 25-FEB-1999; 99US-0121825.  
 PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126264.  
 PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.  
 PR 06-APR-1999; 99US-0128234.  
 PR 08-APR-1999; 99US-0128714.  
 PR 16-APR-1999; 99US-0129845.  
 PR 19-APR-1999; 99US-0130077.  
 PR 21-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.  
 PR 28-APR-1999; 99US-0130891.  
 PR 30-APR-1999; 99US-0131449.  
 PR 30-APR-1999; 99US-0132048.  
 PR 30-APR-1999; 99US-0132407.  
 PR 04-MAY-1999; 99US-0132484.  
 PR 05-MAY-1999; 99US-0132485.  
 PR 06-MAY-1999; 99US-0132486.  
 PR 07-MAY-1999; 99US-0132487.  
 PR 11-MAY-1999; 99US-0132863.  
 PR 14-MAY-1999; 99US-0134256.  
 PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134221.  
 PR 18-MAY-1999; 99US-0134370.  
 PR 19-MAY-1999; 99US-0134768.  
 PR 20-MAY-1999; 99US-0134941.  
 PR 20-MAY-1999; 99US-0135124.  
 PR 21-MAY-1999; 99US-0135353.  
 PR 24-MAY-1999; 99US-0135629.  
 PR 25-MAY-1999; 99US-0136021.  
 PR 27-MAY-1999; 99US-0136392.  
 PR 28-MAY-1999; 99US-0136782.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137502.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 16-JUN-1999; 99US-0139453.  
 PR 17-JUN-1999; 99US-0139492.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.

```
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142330.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144065.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145152.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.

PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 22-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 9.4%; Score 169.6; DB 21; Length 1788;
Best Local Similarity 50.2%; Pred. No. 6.5e-44;
Matches 709; Conservative 0; Mismatches 634; Indels 69; Gaps 9;

Qy 426 cggaaaaacttcccgattccatcaaggagtgatgtgcgcaacggagctaacccacttca 485
    |||||
Db 393 cggcaactcttccactgtcacttaacgcgctttacatccgtaacggtccaaatccacagtt 452
    |||||

Qy 486 cgagcgggtgacaggtcaccaacttcttcgacgagacggtatggttcacgctcaatt 545
    |||||
Db 453 tctccctcggtggtccttaccatctcttcgacggtggtatggttcacgctcaataaaat 512
    |||||

Qy 546 cgaacacggttcagctagctacggttgcgggtttactcagactaacccggtttgttcagga 605
    |||||
Db 513 ccacacggttaagccactctctgtagcagatacgtcaagacttataataacaacgltcga 572
    |||||

Qy 606 acgtcaattgggtcgaccggttttccccaagccatcgggtgagcttcacggccacacgg- 664
    |||||
Db 573 gaaacaaacggagctccggttatgcctaacgtgttttccggattcaacggtgtaacggc 632
    |||||
```

QY 665 --gtattgccgaactcatgctattctacccagagatgcagccggtatagtcgacccgcg 722  
Db 633 gtcagtagctcgtggagctttaacgacagctagggtttaacgacagataaatccgat 692  
QY 723 acacgaaacgggtgtagtaaacccggtttggtctatttcaatggcgggttattggctat 782  
Db 693 taacggcattggttagtaataacaagttagctttcttcagtaacggtctcttgcttt 752  
QY 783 gtcggggatgattttaccttaccaaagtccagatcaatcccaatggagatttataaacct 842  
Db 753 aggtgaatctgattttaccttaccgctcgattaaacgaatcaggagatattgaaacgat 812  
QY 843 tgggtcgggtcggatttgcagacaattagaatccacaatgattgcccacccgaaagtga 902  
Db 813 cggacggtacgatttccgagggaaattagcatgagtagacagctcatctctaaacga 872  
QY 903 cccggaatccggtgaactcttcgctttaagctacgacgtcggttcaaaccttacctaa 962  
Db 873 tccaataacggagaaactttcgcttccggtacggtcggttcca---ccggttttaac 929  
QY 963 atacttccgattctcccgaggaactaaatcacggacgtcgaga---ttcagcttga 1019  
Db 930 atatttccggtttgattccgcccgggaaaaacaaagagacgttccgatattctcgatgac 989  
QY 1020 tcagccaacgatgacgacgatttccgattacagagaacttcgctgcgtacctgacca 1079  
Db 990 gtcctcgtcgtttctccatgacttcgcatcacgaacgtcacgagatttccgagatg 1049  
QY 1080 gcaagtcg-----ttttcaagctcgcgagatgacccggtgggtctcccggtggt 1130  
Db 1050 tcagcttggcatgagatgaaatggtggatttgggtctcgaagggtggtctcccggtgg 1109  
QY 1131 ttacgacaagaacagggtcgaagattcgggattttagacaaatccgccaagattcatc 1190  
Db 1110 tactgataacggaaaaactccaaggcttggagtgattcttaagtagccgagagatgc 1169  
QY 1191 gaacattaagtgattgactccagatgcttccgattgcttccatctctggaacgcttgga 1250  
Db 1170 ggaatgaatggttcgaagtcctggattcattcaatcattcacgctattaatgcttgga 1229  
QY 1251 agagccagaacagatgaagtcggtcgtgatagggtcctgtatgactccaccagactcaat 1310  
Db 1230 tgaagatgatgaaacagcgtcgtttgattgcaccgaatattatgtcattggaacatac 1289  
QY 1311 ttcaacgagctgcagagaatctcaagagtgctcgtgtgataatccgctgaactcaaa 1370  
Db 1290 tttagagagatggatctgttcc---atgcttgggtggagaagggaagatcgatcgt 1346  
QY 1371 aaccggtgaatacaactcgcgtccgatactctccaacgaagatcaacagtcacaactcga 1430  
Db 1347 caccgggatttgagacgtcatcctcagcgagg-----aatctcga 1391  
QY 1431 agcaggatgggtcaacagaaactcgcggcgttaaaacaaattcgtctacttggcttt 1490  
Db 1392 ttctcgtgattatccggcgtttctcgggagatgtagcaggtacgtttacgcggcgat 1451  
QY 1491 agccagcgtggtcctaaagtctcaggttcgtctaaagtgtactcactactcggagaagt 1550  
Db 1452 tggagatccgatgcgaagatcccggtggtggaagctgtgcttaaggagatcg 1511  
QY 1551 taagaatac-----ctttacggcgataaccgttaccgagggagagccctct 1595  
Db 1512 ggaatgattgacggtggcccgatgagatgacggtttcaggttttaccggcggaacccgt 1571  
QY 1596 gttctctcccgaga-----aggagagagaagacgaagatacatctctcg 1643  
Db 1572 ttctcagtagaggtacctggttaaccggagcgaggaggtatggtgtatggtggtgac 1631  
QY 1644 ttctgttcacgacgagaaacatggaatcgaggttcagatagtttaacgctgttagctt 1703  
Db 1632 gtagttcacgatgaagtagcggagaatcgaagtttctggtcgagcgctaaatcgcc 1691  
QY 1704 agaggttgaa-----gcaacgggttaaaacttcgtaacgggttccgtagcggatttcacgg 1757

Db 1692 ggagcttgaataatcgtcgcgcgctgaggttcgcggaagggttccgtagcggattccatg 1751  
QY 1758 tacattcatcggagcgcgatgatttggcgaagc 1789  
Db 1752 gttattgtcaaggaaagtgaaccttaataagc 1783  
RESULT 12  
AAD09395  
ID AAD09395 standard; cDNA; 1788 BP.  
XX  
AC AAD09395;  
XX  
DT 10-SEP-2001 (first entry)  
XX  
DE Arabidopsis thaliana neoxanthin cleavage enzyme, AtNCED2 cDNA.  
XX  
KW Neoxanthin cleavage enzyme; AtNCED2; abscisic acid; ABA; herbicide;  
KW stress tolerance; transgenic plant; plant breeding; antisense-therapy;  
KW plant growth protectant; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..1788  
FT /\*tag= a  
FT /product= "Arabidopsis thaliana AtNCED2 protein"  
XX  
EP1116794-A2.  
XX  
PD 18-JUL-2001.  
XX  
XX 11-JAN-2001; 2001EP-0300218.  
XX  
PR 13-JAN-2000; 2000JP-0010055.  
PR 11-JAN-2001; 2001JP-0003476.  
XX  
PA (RIKE ) RIKEN KK.  
XX  
PI Iuchi S, Kobayashi M, Shinozaki K;  
XX  
XX WPI; 2001-400081/43.  
DR P-PSDB; AAE04783.  
XX  
PT A DNA encoding a protein with a neoxanthin cleavage activity for  
PT producing transgenic plants with improved or decreased stress tolerance  
XX  
XX Example 10; Page 25-29; 101pp; English.  
XX  
CC The invention relates to neoxanthin cleavage enzymes and their  
CC corresponding cDNA molecules. Neoxanthin cleavage enzyme plays a key  
CC role in endogenous abscisic acid (ABA) biosynthesis under drought stress.  
CC Neoxanthin cleavage enzyme is used for improving stress tolerance in a  
CC plant when expressed in a plant cell. The invention also relates to  
CC methods for increasing or decreasing stress tolerance in a plant by  
CC introducing the DNA into the plant, and a transgenic plant into which a  
CC neoxanthin cleavage enzyme is introduced. The improvement of stress  
CC tolerance in plants is useful, for example in plant breeding. Neoxanthin  
CC cleavage enzyme genes are useful for producing transgenic plants. An arid  
CC land can be improved by growing transformant weed for several years and  
CC then removing the weed by specifically lowering stress tolerance in the  
CC weed by inducing an inducible promoter. The present cDNA sequence encodes  
CC Arabidopsis thaliana neoxanthin cleavage enzyme, AtNCED2 protein.  
CC The AtNCED2 cDNA is obtained from an Arabidopsis plant-derived cDNA  
CC library using a cDNA of the CPRD65 (CowPea Responsive to Dehydration)  
CC gene isolated from cowpea plant as a probe.  
XX  
SQ Sequence 1788 BP; 444 A; 444 C; 418 G; 482 T; 0 other;

Query Match

9.4%; Score 169.6; DB 22; Length 1788;

```
Best Local Similarity 50.2%; Pred. No. 6.5e-44;
Matches 709; Conservative 0; Mismatches 634; Indels 69; Gaps 9;

Qy 426 cggaaacttccgattccatcaagaaggtgtatgtgcgaacgagagtaaccacttca 485
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 393 cggacttctccactgtcacttaacggcgcttaacgcgtacgcgtcaaatccacagt 452
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 486 cgaacgggtgacaggtgacacactcttcgacggagacggtatggttcacgcgcgtcaatt 545
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 453 tctccctcgtggttcctaccatctcttcgacggcgaggtatggttcacgcgcataaaaat 512
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 546 cgaacaggttcagtagtactgctgctggttttactcagactaacgggtttgttcagga 605
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 513 ccacaaaggtaaagccactctctgacagatacgtcaagacttataatacaacgctga 572
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 606 acgtcaattgggtcgacgggttttcccaaaagccatcggtgagcttcacggccacacgg- 664
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 573 gaaacaaacggagctccggttatgcttaacgtgttttccggattcaacgggtgaacggc 632
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 665 --gtattcccgactcatgctattctacccagagctgcacgggtatagtcgacccggc 722
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 633 gtcagtgcgtcgagctttaaacggagctaggttttaacccggacagtataatccgggt 692
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 723 acacgaacccggtgtagctaacgcgggtttgtctatttcaatggccggttatgtggttat 782
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 693 taacggcaattggttagctataacagctagctgttcttcagtaacgcgtctcttggttt 752
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 783 gtcggaggtgatttaccttaccaggttcagatcactcccaatggagatttaaaaaacgt 842
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 753 aggtgaatctgatttacctacgcgcgtccgatttaacgaatcaggagatttgaacgat 812
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 843 tggctggttcgatttggatgacataagataatccacatgattcccccacgaagtcga 902
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 813 cggacggtacgatttcgacgggaatagcgtatgagtagatgacgctcatctctaaaccca 872
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 903 cccggaaatccggtgaactcttcggttaagctacgaogctgtttcaaaagccttacctaaa 962
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 873 tccataaacggagaaacttcggttccggtacggtccggttcca---cgtttttaac 929
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 963 atacttcogattctcacggcagcgaactaaataccggagctgcaga---ttcagcttga 1019
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 930 atattccggttgatttcgcgcgggaaacaaagacggttcggatatttcgatgac 989
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1020 tcacccaacgtagatgacgatttcgggattcggattacagaaacttcgctcgtaactgacca 1079
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 990 gtcctccgctgttctccatgacttcggatcacgaaacgtcacgcgatttccgacagat 1049
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1080 gcaagtcg-----ttttcaagctccggagatgacccggtgggttcctccggtggt 1130
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1050 tcagcttggeatgaggatgaacatgttggatttgggttcgcgaaggtggttcctccggttg 1109
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1131 ttacgacaagaacaggttcgcaagattcggatttttagacaaataccgccgaagattcatc 1190
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1110 tactgataacggaaacacccaagcttgagtgatttctcaagtaacgcgcggagatgagtc 1169
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1191 gaacattaaagtggattgattgcagatggttcttcgcttccttcggaacgcttggga 1250
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1170 ggaatgaattggttcgaagttcctggattcaatatcattcacgctattaatgcttggga 1229
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1251 agaccagaacacagatgaagtcgctggtataggttcctcgtagctccaccagactcaat 1310
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1230 tgaagatgatagaacacgcgcgtttgattgcaccgaatattatgctgattgaacatac 1289
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1311 tttaacagagctgcagagaaatcagagtgctcgtctgtgaaatccgcctgaatctcaa 1370
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1290 tttagagaggtatgactcgtggttc---atgctttgggtggagaaggtgaagatcgatcgt 1346
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1371 aacgggaaatcaactccgcctcgatcatctccaacgaagatacaacagtcacacccga 1430
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1347 caccgggatttgagacgtcatcgatcgcagcagg-----aatctcga 1391
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1431 agcagggtggtcaacagaacacgtcctcggcgtaaaaccaaattcgcttacttggcttt 1490
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

XX This invention relates to defence-related signalling genes isolated from  
 CC the sunflower (*Helianthus annuus*). The genes encode a neoxanthin cleavage  
 CC enzyme (NCE), an amino acid permease (AAP) and a glutamic acid rich  
 CC protein (GRP). The signalling gene is useful for increasing the  
 CC resistance of a plant to a pathogen such as fungus, virus, bacterium,  
 CC nematode or insect (e.g. European corn borer), preferably incorporating a  
 CC *Sclerotinia* spp., *Phoma* spp., or *Phomopsis* spp., by stably incorporating a  
 CC construct containing the gene into the genome of the plant. The gene is  
 CC useful for regulating gene expression in a plant, in response to a  
 CC stimulus such as infection with a pathogen, damage from a pathogen,  
 CC hydrogen peroxide, jasmonic acid, methyl jasmonate, salicylic acid,  
 CC oxalic acid or expression of a gene encoding oxalic acid oxidase. The  
 CC genes are also useful for stem-preferred regulation of gene expression in  
 CC a plant. The genes are useful in agriculture, particularly in the  
 CC breeding of crop plants with improved agronomic traits, for modifying  
 CC abscisic acid (ABA) metabolism and for modifying amino acid transport and  
 CC content in plants. The present sequence represents cDNA encoding the  
 CC sunflower neoxanthin cleavage enzyme (NCE).

SQ Sequence 1950 BP; 535 A; 456 C; 453 G; 506 T; 0 other;

Query Match 9.4%; Score 169.2; DB 22; Length 1950;

Best Local Similarity 51.0%; Pred. No. 9.2e-44;

Mismatches 480; Conservative 0; Mismatches 453; Indels 9; Gaps 3;

QY 347 cggctgattctggttcagatcgccggaatttggctccggtgaatgaacagccgctcc 406  
 Db |||||  
 QY 308 cagtgatcccaaacagctttgtctgataattttccacgggagcaactccctccga 367  
 Db |||||  
 QY 407 ggcgtaatctccgggtgcggaacttcccgattccatcaaggagtgatgacga 466  
 Db |||||  
 QY 368 ctgactgtgaagtcatcgaggccactcgaagtgtgcttgcgtgacttccgta 427  
 Db |||||  
 QY 467 acggagctaaccaacttcagagcgggtgacaggtcaccactttccagagacgga 526  
 Db |||||  
 QY 428 atggtccgaaacccaattccctccgcgaggaccctaccactctctcgatggca 487  
 Db |||||  
 QY 527 tgggtcaecgctcaaatccgaacgggttcagctagctacgcttgcgggtttactcga 586  
 Db |||||  
 QY 488 tgcctcagctattcgatctccatggaagagcttcgtatgtagcgatatacaaaa 547  
 Db |||||  
 QY 587 ctacacgggttattcaggagacgtcaattgggtcgaccgggttttcccaagccatcggtg 646  
 Db |||||  
 QY 548 catacaaatattcaatagaaagagggattcccatattcctcaacggttttcag 607  
 Db |||||  
 QY 647 agcttcagggccacacgggtattgcccagctcatgctattc---tacgccagagctgag 703  
 Db |||||  
 QY 608 ggttaaatggtgactgctctgagctcgatgagctcagctcagctcggcgatttttg 667  
 Db |||||  
 QY 704 ccggtatagtcagccggcacacggaacgggtgtagtaaacccggtttgtctatttca 763  
 Db |||||  
 QY 668 ctggacaatttgaccccccaaaaggtattggtcctagccaataccagctcggccttttg 727  
 Db |||||  
 QY 764 atggccggttattggtctatgctggaggatgatttaccattaccaagctcagatcccca 823  
 Db |||||  
 QY 728 gcaacagacttttgcctcgggagatcgatctcccatatcgctcaactagcggccg 787  
 Db |||||  
 QY 824 atgggatttaaaacccgttgctggttgcgatttttgatggacaattagatcccaatga 883  
 Db |||||  
 QY 788 acggtgacatagtcacccgctcggagctgaggaacttcgacggcaactattcatgacatga 847  
 Db |||||  
 QY 884 tggccacccgaaatcgacccgggaatccggtgaactctcgtttaaactcagcgtcg 943  
 Db |||||  
 QY 848 ccgctcaccacaaaatccgatccagtaaacgaaagagcttttcccttcggtccgctcag 907  
 Db |||||  
 QY 944 ttccaagccttacataaactctccgattctccagcagcaggaactaaatcaccggagc 1003  
 Db |||||  
 QY 908 tcccc---ccttctcaacttttccgttttcaacgaaacggagaaaaaacgagcgatg 964  
 Db |||||  
 QY 1004 tggagattcagc---ttgatcgcaacagatgatgcagcagatttccgattacagagaact 1060  
 Db |||||

Db 965 tcccgatcttccaatgacaagcccgctgcttcttccacgacttcgccatccacaaaaact 1024  
 QY 1061 tcgtcgtcgtactgaccagcaagtcggttttcaagctgcggagatgattcccggtgggt 1120  
 Db |||||  
 Db 1025 acgcgatttttcccgatcccaatcggaatgagcccaatgagatgctgggtggggat 1084  
 QY 1121 ctccggttgggtttacgacgaagaacaggtgcgaagatttcgggatttttagacaataacgcg 1180  
 Db |||||  
 Db 1085 ccccggttagcgcgagcgtggaaggtgctcgcggtcggttgcctcctcggtacgga 1144  
 QY 1181 agatttcgaacacattaagtggattgattgctccagattgcttgcctccatctcga 1240  
 Db |||||  
 Db 1145 aagacgagcccgagatgagtggttgggttccgggttttaattgatacatatgcata 1204  
 QY 1241 acgcttgggaagagccgaacacagatgaagtgcgtcgtgtag 1282  
 Db |||||  
 Db 1205 atgcatggagagatggcggagatacgggtggtgatggtgg 1246

RESULT 14  
 AAC57157  
 ID AAC57157 standard; DNA; 491 BP.  
 XX  
 AC AAC57157;  
 XX  
 DT 25-JAN-2001 (first entry)  
 XX  
 DE Pinus radiata transcription factor DNA sequence #603.  
 XX  
 KW Plant; transcription factor; gene expression; eucalyptus; pine; acacia;  
 KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;  
 KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;  
 KW homeodomain zipper; LIM domain; AP2; ERBs; zinc finger domain;  
 KW type 2 Cys2His2; CCAAT box element; MYB; ss.  
 XX  
 OS Pinus radiata.  
 XX  
 PN WO200053724-A2.  
 XX  
 PD 14-SEP-2000.  
 XX  
 PF 09-MAR-2000; 2000WO-US06112.  
 XX  
 PR 11-MAR-1999; 99US-0266513.  
 PR 18-AUG-1999; 99US-0149485.  
 XX  
 PA (GENE-) GENESIS RES & DEV CORP LTD.  
 PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.  
 XX  
 PI Wood M, McGrath A, Shenk MA, Glenn M;  
 XX WPI; 2000-579369/54.  
 XX  
 DR New isolated polynucleotide encoding a plant transcription factor for  
 PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,  
 PT having modified gene expression or modified activity of a polypeptide  
 PT  
 XX  
 PS Claim 1; Page 598; 747pp; English.  
 XX  
 CC The present invention relates to novel plant transcription factors from  
 CC Eucalyptus grandis or Pinus radiata. The present sequence is the coding  
 CC sequence for one such transcription factor. The transcription factor may  
 CC be used to produce a plant having modified gene expression such as a  
 CC woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or  
 CC mahogany species or to modify the activity of a polypeptide in a plant.  
 CC The transcription factors of the present invention are members from the  
 CC following families of regulatory proteins: bZIP, bZIP family of G-box  
 CC binding factors, basic helix-loop-helix zipper,  
 CC homeotic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2  
 CC and ERBs, zinc finger domains of type 2 Cys2His2, CCAAT box elements  
 CC and MYB.  
 XX



SQ Sequence 491 BP; 86 A; 143 C; 171 G; 91 T; 0 other;

Query Match 8.3%; Score 149.4; DB 21; Length 491;  
Best Local Similarity 59.2%; Pred. No. 1.1e-37;  
Matches 255; Conservative 0; Mismatches 176; Indels 0; Gaps 0;

Qy 436 cccgattccatcaaaaggagtgatgtgcgcaacgagagtaaccacttcacgagccggtg 495  
|||  
Db 1 ccggagtgcttagatggaggttaacgtccgcaatggcgcgaatccccgtttcaaacccgc 60  
Qy 496 acaggtccaccacttttcgacggagacggtatggttcacgcccgtcaaatctgaacacggt 555  
|||  
Db 61 ggcggccaccattatttgacggcgatggaatgatacatgcggtgacgtgagacacggg 120  
Qy 556 tcagtagctacgttcggtttactcagactaacgggtttgttcagaaacgtcaattg 615  
|||  
Db 121 aaggttagttacagttgcggttcacggagacggaaggctcgtagcgaggagcggtg 180  
Qy 616 ggtcgacgggtttcccaaacgacatcggtgagcttcacggccacacggtattgccga 675  
|||  
Db 181 gggcgaggctttaccgaagcccatcgggcaactccacggccacggcggtggtgcgc 240  
Qy 676 ctcatgctattctacgcagagctgcagcgggtatagtcgacccgggcacacggacccgt 735  
|||  
Db 241 ctgctcgtcgtatggtgcggcggtctgcgggtcgttcacacgggaaggatgggc 300  
Qy 736 gtatgaacgcgggtttggtctatttcgaatggcggttattgggtatgtcggaggtgat 795  
|||  
Db 301 gtggttaatgcgggtggcttttttaacggcgctcgtcgtatgtccgaagacgat 360  
Qy 796 ttaccttaccagttcagatcacctcccgaatggagatttaaaacggttgcgttcgat 855  
|||  
Db 361 ctcccgtagcgtcagggtagcgggtgacgggtgacggcgatctggtgacgacgggcaggttcgat 420  
Qy 856 ttgatggaca 866  
|||  
Db 421 ttcgacgggca 431

## RESULT 15

AAC57162  
ID AAC57162 standard; DNA; 386 BP.

XX AC AAC57162;

XX DT 25-JAN-2001 (first entry)

XX DE Pinus radiata transcription factor DNA sequence #608.

XX KW Plant; transcription factor; gene expression; eucalyptus; pine; acacia;  
KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;  
KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;  
KW homeodomain zipper; LIM domain; AP2; EREBs; zinc finger domain;  
KW type 2 Cys2His2; CCAAT box element; MYB; ss.

XX OS Pinus radiata.

XX FN WO200053724-A2.

XX PD 14-SEP-2000.

XX PF 09-MAR-2000; 2000WO-US06112.

XX PR 11-MAR-1999; 99US-0266513.

XX PR 18-AUG-1999; 99US-0149485.

XX PA (GENE-) GENESIS RES & DEV CORP LTD.

XX PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.

XX PI Wood M, McGrath A, Shenk MA, Glenn M;

XX DR WPI; 2000-579369/54.

XX

PT New isolated polynucleotide encoding a plant transcription factor for  
PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,  
PT having modified gene expression or modified activity of a polypeptide

XX

PS Claim 1; Page 599; 747pp; English.

XX

CC The present invention relates to novel plant transcription factors from  
CC Eucalyptus grandis or Pinus radiata. The present sequence is the coding  
CC sequence for one such transcription factor. The transcription factor may  
CC be used to produce a plant having modified gene expression such as a  
CC woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or  
CC mahogany species or to modify the activity of a polypeptide in a plant.  
CC The transcription factors of the present invention are members from the  
CC following families of regulatory proteins: bZIP, bZIP family of G-box  
CC binding factors, basic helix-loop-helix zipper,  
CC homeotic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2  
CC and EREBs, zinc finger domains of type 2 Cys2His2, CCAAT box elements  
CC and MYB.

XX SQ Sequence 386 BP; 89 A; 105 C; 106 G; 86 T; 0 other;

Query Match 8.0%; Score 143.4; DB 21; Length 386;

Best Local Similarity 70.0%; Pred. No. 8.4e-36;

Matches 208; Conservative 0; Mismatches 86; Indels 3; Gaps 1;

Qy 1103 agatgattccggtgggtctccggtgtttacgacagaagaacaggtcgcaagattcggga 1162  
|||  
Db 1 aaatgatcagaggcggttctccagttattcacacaagaagaaggtcccgctcgggc 60

Qy 1163 ttttagcaataacgcggaagatttcacgaacattaaagtggattgactccagattgct 1222  
|||  
Db 61 ttctgcccaaatatgctctgacgagagtgagctgaaatggatcgaggtcccgattgct 120

Qy 1223 tctgcttccatctctggaacgcttgggaagagccagaaacagatgaagtctgctgatag 1282  
|||  
Db 121 tctgcttccatctctggaacgcttgggaagagccagaaacagatgaagtctgctgatag 177

Qy 1283 ggtcctgtagactccaccagactcaattttcaacgagctgacgagaatcttcaagagtg 1342  
|||  
Db 178 gctcctgtagactccaccgagcgccttcaacgagctgacgagaatctgacgagcgtcgaggagt 237

Qy 1343 tctgtctgaaatccgctgaatctcaaacccggtgaatcaactcgcgtccgatca 1399  
|||  
Db 238 ttctgtcgaaatccgctgaatctcaaacccggttgtccaccagacgagatca 294

Search completed: July 24, 2002, 06:25:34

Job time: 6890 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 24, 2002, 03:54:00 ; Search time 57.66 Seconds  
(without alignments)  
7668.059 Million cell updates/sec

Title: US-09-758-269-5  
Perfect score: 1800  
Sequence: 1 atggcttttttcagcgaac.....tggcgaagcaggctgtgtga 1800

Scoring table:  
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38.8	2.2	7218	1 US-08-232-463-14	Sequence 14, Appl
2	34.8	1.9	4695	6 5225348-3	Patent No. 5225348
3	34.2	1.9	1221	1 US-08-434-881-1	Sequence 1, Appl
4	34.2	1.9	1221	3 US-08-977-771-1	Sequence 1, Appl
5	34.2	1.9	1221	4 US-09-361-773-1	Sequence 1, Appl
6	34.2	1.9	1370	4 US-09-026-408-12	Sequence 12, Appl
7	34.2	1.9	1371	4 US-09-026-408-1	Sequence 1, Appl
8	33.4	1.9	788	4 US-08-991-789A-177	Sequence 177, App
9	33.4	1.9	788	4 US-09-062-451-177	Sequence 177, App
10	33.4	1.9	1047	5 PCT-US96-10986-11	Sequence 11, Appl
11	33.4	1.9	1753	6 5225348-2	Sequence 11, Appl
12	33.4	1.9	3680	2 US-08-494-907-1	Patent No. 5225348
13	33.4	1.9	2680	5 PCT-US96-10986-1	Sequence 1, Appl
14	33.4	1.9	5076	2 US-08-494-907-2	Sequence 2, Appl
15	33.4	1.9	5076	5 PCT-US96-10986-2	Sequence 2, Appl
16	33.4	1.9	6170	2 US-08-494-907-4	Sequence 4, Appl
17	33.4	1.9	6170	5 PCT-US96-10986-4	Sequence 4, Appl
18	33.4	1.9	6387	2 US-08-494-907-3	Sequence 3, Appl
19	33.4	1.9	6387	5 PCT-US96-10986-3	Sequence 3, Appl
20	33.4	1.9	7198	4 US-08-994-035C-4	Sequence 4, Appl
21	33.4	1.9	4411529	4 US-09-103-840A-1	Sequence 1, Appl
22	33.4	1.9	607	4 US-09-385-982-525	Sequence 525, App
23	33	1.8	2128	2 US-08-371-377-16	Sequence 16, Appl
24	33	1.8	80161	3 US-09-036-987A-1	Sequence 1, Appl
25	33	1.8	80161	4 US-09-370-700-1	Sequence 1, Appl
26	33	1.8	1092	4 US-09-077-675A-15	Sequence 15, Appl
27	32.6				

Sequence 14, Appl  
Sequence 14, Appl  
Sequence 33, Appl  
Sequence 33, Appl  
Sequence 33, Appl  
Sequence 33, Appl  
Sequence 33, Appl  
Sequence 1, Appl  
Sequence 1, Appl  
Sequence 25, Appl  
Sequence 1, Appl  
Sequence 1, Appl  
Sequence 17, Appl  
Sequence 17, Appl  
Sequence 448, App  
Sequence 14, Appl  
Sequence 12, Appl  
Sequence 12, Appl

ALIGNMENTS

RESULT 1  
US-08-232-463-14  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHEIFLINGER, F.  
; APPLICANT: FALKNER, F. G.  
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,463  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/935,313  
; FILING DATE:  
; APPLICATION NUMBER: EP 91 114 300.6  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 30472/114 IMM  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)836-9300  
; TELEFAX: (703)683-4109  
; TELEX: 899149  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7218 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; CLONE: pTZgpt-F1s  
US-08-232-463-14

```
Query Match          2.2%; Score 38.8; DB 1; Length 7218;
Best Local Similarity 7.9%; Pred. No. 0.049;
Matches 16; Conservative 112; Mismatches 74; Indels 0; Gaps 0;

Qy 21 ggctgagggtttctggagatgcttgggtggcaatcatactcagccgcatatcgcttctc 80
Db 1041 GGCTGAGGTCAGGAGCTTGGCATYVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVV 1100

Qy 81 tcaagctccgacttgagttatttagctctccttaccatgcccagtcgtgcacacgtaa 140
Db 1101 YVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVV 1160

Qy 141 gctcaatgttcacatcgcttcacactccctccagctcttcattcccttaagcaatc 200
Db 1161 YVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVV 1220

Qy 201 aaactctcccgccattgtgtt 222
Db 1221 YVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVVV 1242

RESULT 2
5225348-3
;PATENT NO. 5225348
;APPLICANT: HAGATA, SHIGEKAZU;SUGANO, SUMIO;KIM, DONG W.;
;UETSUKI, TAICHI;KAZIRO, YOSHITO
;TITLE OF INVENTION: DNA FRAGMENT AND EXPRESSION PLASMID
;CONTAINING THE DNA FRAGMENT
;NUMBER OF SEQUENCES: 9
;CURRENT APPLICATION DATA:
;APPLICATION NUMBER: US/07/447,823
;FILING DATE: 08-DEC-1989
;SEQ ID NO:3:
;LENGTH: 4695
5225348-3

Query Match          1.9%; Score 34.8; DB 6; Length 4695;
Best Local Similarity 48.1%; Pred. No. 0.78;
Matches 99; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

Qy 340 cctaaacgctgattcctcagtggttcagatcgccggaattttgtcccggtgaatgaacag 399
Db 3386 cctgattggattgccacacggctcacattgcagcaagtttgctgagctgaaggaaag 3445

Qy 400 cccgtccggcgttaattcttcggtggtgcggaacttccgattccatcaaaaggagtgat 459
Db 3446 attgatcgctgttcgttaaaagctgaagatggcctaaattcttgagctcgtgat 3505

Qy 460 gtgcgcaacgagctaacccacttcacagcgggtgacaggtcaccactttctcgacgga 519
Db 3506 gctgcattgtgatgtgttcctggcagcccatgtgtgtgagagcttctcagactat 3565

Qy 520 gacggtatggttcacgcgcgtcaaat 545
Db 3566 ccacctttgggttaaggatgactactt 3591

RESULT 3
US-08-434-881-1/c
; Sequence 1, Application US/08434881
; Patent No. 5804376
; GENERAL INFORMATION:
; APPLICANT: Braxton, Scott M.
; APPLICANT: Wilde, Craig G.
; APPLICANT: Diep, Dinh
; TITLE OF INVENTION: Pancreas-Derived Serpin
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3330 Hillview Avenue
; CITY: Palo Alto
; STATE: California
```

```
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,881
; FILING DATE: Herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Luther, Barbara J.
; REGISTRATION NUMBER: 33954
; REFERENCE/DOCKET NUMBER: PF0035 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-852-0195
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1221 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; IMMEDIATE SOURCE:
; LIBRARY: Pancreas
; CLONE: 222689
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1221
; US-08-434-881-1

Query Match          1.9%; Score 34.2; DB 1; Length 1221;
Best Local Similarity 54.3%; Pred. No. 0.51;
Matches 69; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Qy 97 agttattgtagctcttaccatggtccagtcggtgcacgtaagctcaatgtttctatct 156
Db 788 ACTTCTTCTATATCCATACCTTCTGCAGGAAGTATGATAATTAAAGCTAAATTCATCACC 729

Qy 157 gcgcttcacatctccagctcttcttccctaaagcaatcctcaactctcccgccatt 216
Db 728 TTGTAAGACAACTTAAACTTGTTAGGAGAGATTGAGAAATAACCATATATTTT 669

Qy 217 gttgta 223
Db 668 GTTCTCA 662

RESULT 4
US-08-977-771-1/c
; Sequence 1, Application US/08977771
; Patent No. 6013448
; GENERAL INFORMATION:
; APPLICANT: Braxton, Scott M.
; APPLICANT: Wilde, Craig G.
; APPLICANT: Diep, Dinh
; TITLE OF INVENTION: Pancreas-Derived Serpin
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3330 Hillview Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/977,771  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/ 434,881  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Luther, Barbara J.  
; REGISTRATION NUMBER: 33954  
; REFERENCE/DOCKET NUMBER: PF0035 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-852-0195  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1221 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; IMMEDIATE SOURCE:  
; LIBRARY: Pancreas  
; CLONE: 222689  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..1221  
US-08-977-771-1

Query Match 1.9%; Score 34.2; DB 3; Length 1221;  
Best Local Similarity 54.3%; Pred. No. 0.51;  
Matches 69; Conservative 0; Mismatches 58; Indels 0; Gaps 0;  
Qy 97 agttattgagctcttaccatgcccagctcgtgcacacgtaagctcaatgtttcatct 156  
Db 788 ACTTCTTATATCCATACCTCTGCGAGGAGTATGATTAATTAAGCTAAATTCACCT 729  
Qy 157 gcgcttcacactcctccagctcttctccctcaagcaatcatcaactctcccgccatt 216  
Db 728 TTGTAAGACAATCTTAAACTTGGTAGTTCAGGAGAGATTTCAGAAAAATAACCATATTTT 669  
Qy 217 gttgta 223  
Db 668 GTTCTCA.662

RESULT 5  
US-09-361-773-1/c  
; Sequence 1, Application US/09361773  
; Patent No. 6197519  
; GENERAL INFORMATION:  
; APPLICANT: Braxton, Scott M.  
; APPLICANT: Wilde, Craig G.  
; APPLICANT: Diep, Dinh  
; TITLE OF INVENTION: Pancreas-Derived Serpin  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3330 Hillview Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/361,773  
; FILING DATE:  
; CLASSIFICATION:

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/977,771  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/434,881  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Luther, Barbara J.  
; REGISTRATION NUMBER: 33954  
; REFERENCE/DOCKET NUMBER: PF0035 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-852-0195  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1221 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; IMMEDIATE SOURCE:  
; LIBRARY: Pancreas  
; CLONE: 222689  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..1221  
US-09-361-773-1  
  
Query Match 1.9%; Score 34.2; DB 4; Length 1221;  
Best Local Similarity 54.3%; Pred. No. 0.51;  
Matches 69; Conservative 0; Mismatches 58; Indels 0; Gaps 0;  
Qy 97 agttattgagctcttaccatgcccagctcgtgcacacgtaagctcaatgtttcatct 156  
Db 788 ACTTCTTATATCCATACCTCTGCGAGGAGTATGATTAATTAAGCTAAATTCACCT 729  
Qy 157 gcgcttcacactcctccagctcttctccctcaagcaatcatcaactctcccgccatt 216  
Db 728 TTGTAAGACAATCTTAAACTTGGTAGTTCAGGAGAGATTTCAGAAAAATAACCATATTTT 669  
Qy 217 gttgta 223  
Db 668 GTTCTCA.662  
  
RESULT 6  
US-09-026-408-12/c  
; Sequence 12, Application US/09026408  
; Patent No. 6303338  
; GENERAL INFORMATION:  
; APPLICANT: Ni et al.  
; TITLE OF INVENTION: PANCREAS-DERIVED PLASMINOGEN ACTIVATOR  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 NEW YORK AVENUE, SUITE 600  
; CITY: WASHINGTON  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/026,408  
; FILING DATE: Herewith  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/934,011  
; FILING DATE: 15-AUG-1997



CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Potter, Jane E. R.  
REGISTRATION NUMBER: 33,332  
REFERENCE/DOCKET NUMBER: 210121.419C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 177:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 788 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 177:  
US-08-991-789A-177

Query Match 1.9%; Score 33.4; DB 4; Length 788;  
Best Local Similarity 48.2%; Pred. No. 0.71;  
Matches 94; Conservative 0; Mismatches 101; Indels 0; Gaps 0;  
QY 340 cctaaacggcgtgctcctagttcagatcgccggaattttgctccggtgaatgaacag 399  
DB 536 CTGTATTGGATTGCCACACGGCTACATTGCTGATGCAAGTTTGTGCTGAGCTGAAGGAAAG 595  
QY 400 cccgtccggcgtgaattcttcgggtggtcggaatactccgattccatcaaaaggagtgtat 459  
DB 596 ATTGATCGCGTTCGTGTAAGGCTGGAAGATGGCCCTAAATTTCTTGAAGTCTGGTGAT 655  
QY 460 gtgcgaacggagcgttaaccacttcacgacggcggtgacaggtcaccactttcttcacgga 519  
DB 656 GCTGCAATTGTTGATATGTTCTTCTGCAAGCCCATGTGTGTTGAGAGCTTCTCAGACTAT 715  
QY 520 gacggtatggttcac 534  
DB 716 CCACCTTTGGTGGC 730

RESULT 9  
US-09-062-451-177  
Sequence 177, Application US/09062451  
Patent No. 6344550  
GENERAL INFORMATION:  
APPLICANT: Frudakis, Tony N.  
APPLICANT: Smith, John M.  
APPLICANT: Reed, Steven G.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
TREATMENT AND DIAGNOSIS OF BREAST CANCER  
NUMBER OF SEQUENCES: 297  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/062,451  
FILING DATE: 04-APR-1997  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.419C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 177:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 788 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-062-451-177  
Query Match 1.9%; Score 33.4; DB 4; Length 788;  
Best Local Similarity 48.2%; Pred. No. 0.71;  
Matches 94; Conservative 0; Mismatches 101; Indels 0; Gaps 0;  
QY 340 cctaaacggcgtgctcctagttcagatcgccggaattttgctccggtgaatgaacag 399  
DB 536 CTGTATTGGATTGCCACACGGCTACATTGCTGATGCAAGTTTGTGCTGAGCTGAAGGAAAG 595  
QY 400 cccgtccggcgtgaattcttcgggtggtcggaatactccgattccatcaaaaggagtgtat 459  
DB 596 ATTGATCGCGTTCGTGTAAGGCTGGAAGATGGCCCTAAATTTCTTGAAGTCTGGTGAT 655  
QY 460 gtgcgaacggagcgttaaccacttcacgacggcggtgacaggtcaccactttcttcacgga 519  
DB 656 GCTGCAATTGTTGATATGTTCTTCTGCAAGCCCATGTGTGTTGAGAGCTTCTCAGACTAT 715  
QY 520 gacggtatggttcac 534  
DB 716 CCACCTTTGGTGGC 730  
RESULT 10  
US-08-494-907-11/c  
Sequence 11, Application US/08494907  
Patent No. 5955298  
GENERAL INFORMATION:  
APPLICANT: Thomasow, Linda S  
APPLICANT: Bangera, Mahalaxmi  
APPLICANT: Weller, David M  
APPLICANT: Cook, R. James  
TITLE OF INVENTION: Sequences for Production of  
2,4-diacetylphloroglucinol and Methods  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Margaret A. Connor, USDA-ARS  
STREET: 800 Buchanan Street  
CITY: Albany  
STATE: CA  
COUNTRY: USA  
ZIP: 94710  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/494,907  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Connor, Margaret A  
REGISTRATION NUMBER: 30043  
REFERENCE/DOCKET NUMBER: 0009.95  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (510) 559-6067  
TELEFAX: (510) 559-5777  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1047 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO

```
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Pseudomonas fluorescens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (1..1047)
; OTHER INFORMATION: /note="ph1D DNA sequence. SEQ ID
; US-08-494-907-11
; OTHER INFORMATION: NO:12 is translation (protein) of SEQ ID NO:11."

Query Match      1.9%; Score 33.4; DB 2; Length 1047;
Best Local Similarity 48.2%; Pred. No. 0.85;
Matches 94; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

Qy 585 gactaacgggtttgttcagggaacgtcaattggtgcgacgggttttccccaaagccatcgg 644
Db 937 GGCTAAACGGATGATTCAAAACACCCCAAGTCAATGAACGGTATCTGGTCTTGCCCATCGA 878

Qy 645 tgagcttcacggccacacggattgcccgaactcatgtattacgccagagctgcagc 704
Db 877 TGAACCTCGGTGCATACCGGCTTTACCCACCGCAGCATCGTGTATGAGCGAGGCTCG 818

Qy 705 cggatagtcgacccggcacacggagcgggtgtgactaacccgggtttggtctatttcaa 764
Db 817 CCGCATGCTTCAATCGCGCGCGCCAGGCCATTGAAAACCGGGCTTGACCCAGGACGA 758

Qy 765 tggccgggtattggc 779
Db 757 CATCCGGATGGTGC 743

RESULT 11
PCT-US96-10986-11/c
; Sequence 11, Application PC/TUS9610986
; GENERAL INFORMATION:
; TITLE OF INVENTION: Sequences for Production of
; 2,4-Diacetylphloroglucinol and Methods
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Stephan A. Pendorf, DOMINIK & STEIN
; STREET: 600 N. West Shore Boulevard, Suite 1000
; CITY: Tampa
; STATE: FL
; COUNTRY: USA
; ZIP: 33609
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/10986
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pendorf, Stephan A.
; REGISTRATION NUMBER: 32665
; REFERENCE/DOCKET NUMBER: A700.320
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (813) 289-2966
; TELEFAX: (813) 289-2967
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1047 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Pseudomonas fluorescens
```

```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (1..1047)
; OTHER INFORMATION: /note="ph1D DNA sequence. SEQ ID
; PCT-US96-10986-11
; OTHER INFORMATION: NO:12 is translation (protein) of SEQ ID NO:11."

Query Match      1.9%; Score 33.4; DB 5; Length 1047;
Best Local Similarity 48.2%; Pred. No. 0.85;
Matches 94; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

Qy 585 gactaacgggtttgttcagggaacgtcaattggtgcgacgggttttccccaaagccatcgg 644
Db 937 GGCTAAACGGATGATTCAAAACACCCCAAGTCAATGAACGGTATCTGGTCTTGCCCATCGA 878

Qy 645 tgagcttcacggccacacggattgcccgaactcatgtattacgccagagctgcagc 704
Db 877 TGAACCTCGGTGCATACCGGCTTTACCCACCGCAGCATCGTGTATGAGCGAGGCTCG 818

Qy 705 cggatagtcgacccggcacacggagcgggtgtgactaacccgggtttggtctatttcaa 764
Db 817 CCGCATGCTTCAATCGCGCGCGCCAGGCCATTGAAAACCGGGCTTGACCCAGGACGA 758

Qy 765 tggccgggtattggc 779
Db 757 CATCCGGATGGTGC 743

RESULT 12
5225348-2
; Patent No. 5225348
; APPLICANT: HAGATA, SHIGEKAZU; SUGANO, SUMIO; KIM, DONG W.;
; UETSUKI, TAICHI; KAZIRO, YOSHITO
; TITLE OF INVENTION: DNA FRAGMENT AND EXPRESSION PLASMID
; CONTAINING THE DNA FRAGMENT
; NUMBER OF SEQUENCES: 9
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/447,823
; FILING DATE: 08-DEC-1989
; SEQ ID NO:2:
; LENGTH: 1753
5225348-2

Query Match      1.9%; Score 33.4; DB 6; Length 1753;
Best Local Similarity 48.2%; Pred. No. 1.2;
Matches 94; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

Qy 340 cctaaacggctgatcctcctagttctcagatcgccggaaattttgctccggtgaatgaacag 399
Db 1137 cctgattggattgccacacggctccatcgcatgcaagtttgcgagcgaaagaaaag 1196

Qy 400 cccgtccgggtaattcttcgggtggtcggaatactcccgattccatcaaaaggagtgat 459
Db 1197 attgatcgccgttctgttataaaagctggaagatggccctaaattcttgagctcgtgat 1256

Qy 460 gtgcgaacggagctaacccacttcacgacggcggtgcaggtcacccacttttcgacgga 519
Db 1257 gctgccattgtgatgttcttcctcggcaagcccatgtgtgtgagagcttctcagactat 1316

Qy 520 gacggtatggttcac 534
Db 1317 ccaccttgggtcgc 1331

RESULT 13
US-08-494-907-1/c
; Sequence 1, Application US/08494907
; Patent No. 5955298
; GENERAL INFORMATION:
; APPLICANT: Thomasow, Linda S
; APPLICANT: Banger, Mahalaxmi
```



```

; APPLICANT: Weller, David M
; APPLICANT: Cook, R. James
; TITLE OF INVENTION: Sequences for Production of
; TITLE OF INVENTION: 2,4-Diacetylphloroglucinol and Methods
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Margaret A. Connor, USDA-ARS
; STREET: 800 Buchanan Street
; CITY: Albany
; STATE: CA
; COUNTRY: USA
; ZIP: 94710
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/494,907
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Connor, Margaret A
; REGISTRATION NUMBER: 30043
; REFERENCE/DOCKET NUMBER: 0009.95
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 559-6067
; TELEFAX: (510) 559-5777
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORGANISM: Pseudomonas fluorescens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: complement (2889..3680)
; OTHER INFORMATION: /note="phlA, transcribed from
; OTHER INFORMATION: right to left"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: complement (1210..1917)
; OTHER INFORMATION: /note="phlB, transcribed from
; OTHER INFORMATION: right to left"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: complement (1689..2855)
; OTHER INFORMATION: /note="phlC, transcribed from
; OTHER INFORMATION: right to left"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: complement (2..1048)
; OTHER INFORMATION: /note="phlD, transcribed from
; OTHER INFORMATION: right to left"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 2118..3371
; OTHER INFORMATION: /note="phlR, transcribed from left
; OTHER INFORMATION: to right"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..3680
; OTHER INFORMATION: /note="SEQ ID NO:1 contains genes
; OTHER INFORMATION: necessary for Phl synthesis."
; US-08-494-907-1

```

Query Match 1.9%; Score 33.4; DB 2; Length 3680;

```

Best Local Similarity 48.2%; Pred. No. 1.9;
Matches 94; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
QY 585 gactaacgggtttgttcagaaacgtcaattgggtcgaccgggttttccccaaagccatcgg 644
DB 938 GGCTAAACGGATGATTCAAAACACCCCACTCAATGAACGGTATCTGGTCTTGCCCATCGA 879
QY 645 tgagttcacggccacacgggtattgcccagactatgtattctacgccagagctgcagc 704
DB 878 TGAACCTGGGTGCATACCGGCTTTACCCACCGCAGCATCGTGTATGACGAGAGGCTCG 819
QY 705 cggatatgtcagccggcacacgggaacgggtgtagctaacgggtttggtattttaa 764
DB 818 CCGCATGCTTCAATCGCGCGCGCCGATTAACCAACGGCGCTTGACCAACGGACCGA 759
QY 765 tggccgggtattggc 779
DB 758 CATCCGGATGGTCG 744

RESULT 14
PCT-US96-10986-1/c
; Sequence 1, Application PC/TUS9610986
; GENERAL INFORMATION:
; TITLE OF INVENTION: Sequences for Production of
; TITLE OF INVENTION: 2,4-Diacetylphloroglucinol and Methods
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Stephan A. Pendorf, DOMINIK & STEIN
; STREET: 600 N. West Shore Boulevard, Suite 1000
; CITY: Tampa
; STATE: FL
; COUNTRY: USA
; ZIP: 33609
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/10986
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pendorf, Stephan A.
; REGISTRATION NUMBER: 32665
; REFERENCE/DOCKET NUMBER: A700.320
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (813) 289-2966
; TELEFAX: (813) 289-2967
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Pseudomonas fluorescens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: complement (2889..3680)
; OTHER INFORMATION: /note="phlA, transcribed from
; OTHER INFORMATION: right to left"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: complement (1210..1917)
; OTHER INFORMATION: /note="phlB, transcribed from
; OTHER INFORMATION: right to left"
; FEATURE:
; NAME/KEY: misc feature

```

LOCATION: complement (1689..2855)  
OTHER INFORMATION: /note= "phlC, transcribed from  
OTHER INFORMATION: right to left"

FEATURE:

NAME/KEY: misc feature  
LOCATION: complement (2..1048)  
OTHER INFORMATION: /note= "phlD, transcribed from  
OTHER INFORMATION: right to left"

FEATURE:

NAME/KEY: misc feature  
LOCATION: 2118..3371  
OTHER INFORMATION: /note= "phlR, transcribed from left  
OTHER INFORMATION: to right"

FEATURE:

NAME/KEY: misc feature  
LOCATION: 1..3680  
OTHER INFORMATION: /note= "SEQ ID NO:1 contains genes  
OTHER INFORMATION: necessary for Phl synthesis."

PCT-US96-10986-1

Query Match 1.9%; Score 33.4; DB 5; Length 3680;  
Best Local Similarity 48.2%; Pred. No. 1.9;  
Matches 94; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

Qy 585 gactaacccggtttgttcaggaaacgtcaattgggtcgaccggttttccccaaagccatcgg 644

Db 938 GGCATAACGGATGATTCATAAACACACCCAAAGTCAATGAACGGTATCTGGTTCGCCATCGA 879

Qy 645 tgagcttcacggccacacccggtattggccgactatgctattctacgcagagctgcagc 704

Db 878 TGAATTCCGCTGCATACCGGCTTTACCCACGGCAGCATCGTATGAGCGAGAGGCTCG 819

Qy 705 cggctatagtcgaccggccacacgggaacccggtgtagctaaccggtttggtctatttcaa 764

Db 818 CCGCATGTCTTCAATCGCGCGCGCCAGCCATTGAAAACGGGGCTTGACCCAGGACGA 759

Qy 765 tggccggttatggc 779

Db 758 CATCCGGATGGTCGC 744

## RESULT 15

US-08-494-907-2/c

Sequence 2, Application US/08494907

Patent No. 5955298

GENERAL INFORMATION:

APPLICANT: Thomasow, Linda S

APPLICANT: Bangera, Mahalaxmi

APPLICANT: Weller, David M

APPLICANT: Cook, R. James

TITLE OF INVENTION: Sequences for Production of

TITLE OF INVENTION: 2,4-Diacetylphloroglucinol and Methods

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: Margaret A. Connor, USDA-ARS

STREET: 800 Buchanan Street

CITY: Albany

STATE: CA

COUNTRY: USA

ZIP: 94710

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/494,907

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Connor, Margaret A

REGISTRATION NUMBER: 30043

REFERENCE/DOCKET NUMBER: 0009.95  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (510) 559-6067

TELEFAX: (510) 559-5777

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 5076 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Pseudomonas fluorescens

FEATURE:

NAME/KEY: misc feature

LOCATION: complement (4285..5076)

OTHER INFORMATION: /note= "phlA, transcribed from

OTHER INFORMATION: right to left"

FEATURE:

NAME/KEY: misc feature

LOCATION: complement (2606..3313)

OTHER INFORMATION: /note= "phlB, transcribed from

OTHER INFORMATION: right to left"

FEATURE:

NAME/KEY: misc feature

LOCATION: complement (3085..4251)

OTHER INFORMATION: /note= "phlC, transcribed from

OTHER INFORMATION: right to left"

FEATURE:

NAME/KEY: misc feature

LOCATION: complement (1398..2444)

OTHER INFORMATION: /note= "phlD, transcribed from

OTHER INFORMATION: right to left"

FEATURE:

NAME/KEY: misc feature

LOCATION: 3514..4767

OTHER INFORMATION: /note= "phlR, transcribed from left

OTHER INFORMATION: to right"

FEATURE:

NAME/KEY: misc feature

LOCATION: complement (2..1270)

OTHER INFORMATION: /note= "phlE, transcribed from

OTHER INFORMATION: right to left"

FEATURE:

NAME/KEY: misc feature

LOCATION: 1..5076

OTHER INFORMATION: /note= "SEQ ID NO:2 contains genes

OTHER INFORMATION: involved in synthesis, activity, and/or export of

OTHER INFORMATION: Phl."

US-08-494-907-2

Query Match 1.9%; Score 33.4; DB 2; Length 5076;  
Best Local Similarity 48.2%; Pred. No. 2.4;  
Matches 94; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

Qy 585 gactaacccggtttgttcaggaaacgtcaattgggtcgaccggttttccccaaagccatcgg 644

Db 2334 GGCATAACGGATGATTCATAAACACACCCAAAGTCAATGAACGGTATCTGGTTCGCCATCGA 2275

Qy 645 tgagcttcacggccacacccggtattggccgactatgctattctacgcagagctgcagc 704

Db 2274 TGAACCTTCGGGTGCATACCGGCTTTACCCACCGCAGCATCGTGTATGAGCGAGAGGCTCG 2215

Qy 705 cggctatagtcgaccggccacacgggaacccggtgtagctaaccggtttggtctatttcaa 764

Db 2214 CCGCATGTCTTCAATCGCGCGCGCCAGGCCATTGAAAACCGGGCTTGACCCAGGACGA 2155

Qy 765 tggccggttatggc 779

Db 2154 CATCCGGATGGTCGC 2140

Search completed: July 24, 2002, 06:21:27  
Job time: 8847 sec

**This Page Blank (uspto)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 24, 2002, 05:14:15 ; Search time 1694.89 Seconds  
(without alignments)  
14333.980 Million cell updates/sec

Title: US-09-758-269-5

Perfect score: 1800

Sequence: 1 atggcttcttcacggcaac.....tggcgaagcaggtcgtgtga 1800

Scoring table:

IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:\*

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	425.6	23.6	720	10	BM412731
2	417	23.2	805	10	BM408615
3	416.2	23.1	781	12	BH549344
4	382.8	21.3	643	9	AW933324
5	375.6	20.9	787	10	BM408565
6	344	19.1	592	10	BF113346
7	332.4	18.5	566	10	BE461924
8	323.2	18.0	553	10	BM085672
9	318	17.7	559	10	BM085005
10	316.6	17.6	520	10	BE432853
11	313.8	17.4	742	12	BH559796
12	312.8	17.4	594	12	BH458011
13	309.4	17.2	547	10	BI974879
14	308.6	17.1	547	10	BE437072
15	303.4	16.9	617	10	BE458861
16	300.4	16.7	546	10	BM084948
17	295.8	16.4	618	9	AW443298

18	294	16.3	564	10	BM085488
19	292.8	16.3	627	10	BF051297
20	289.2	16.1	509	10	BE451573
21	288.8	16.0	495	10	BE434930
22	276.2	15.3	490	10	BF050563
23	275.4	15.3	657	9	AA556214
24	267.4	14.9	503	12	B27476
25	254.8	14.2	566	10	BE459895
26	249.2	13.8	450	9	AW930245
27	248	13.8	552	10	BG039692
28	245	13.6	778	12	BH579654
29	235.8	13.2	600	12	BH458018
30	233.2	13.0	424	12	BH456327
31	224.4	12.5	337	12	BH579641
32	220.4	12.2	410	10	BE522530
33	217	12.1	446	10	BE471089
34	216.6	12.0	713	10	BI928174
35	209.6	11.6	495	9	AU084510
36	207	11.5	502	12	CNS000KV
37	205.6	11.4	590	10	BG592987
38	203.2	11.3	447	9	AW933245
39	202.8	11.3	492	9	AU084760
40	201.4	11.2	412	9	AV419581
41	195.2	10.8	680	12	AQ969739
42	185.8	10.3	549	10	BF425834
43	173.6	9.6	500	9	AW289745
44	164.6	9.1	678	12	BH593312
45	152.6	8.5	267	10	BE529289

## ALIGNMENTS

### RESULT 1

BM412731

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BM412731 EST587047 tomato breaker fruit Lycopersicon esculentum cDNA clone  
CLEG60N24 5' end, mRNA sequence.

BM412731.1 GI:18264350  
EST.  
tomato.

Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;  
Lycopersicon.

1 (bases 1 to 720)  
Alcala,J., Vrebalov,J., White,R., Vision,T., Karamycheva,S.A., Tsai  
,J., Bougri,O., Kirkness,E., Utterback,T., Van Aken,S., Ronning  
,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.  
Generation of ESTs from tomato fruit tissue, breaker stage (2002)  
Unpublished (2002)  
Contact: CUGI  
Clemson University  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
This clone is available through the Clemson University Genomics  
Institute  
Seq primer: T3.

### FEATURES

source

Location/Qualifiers  
1..720  
/organism="Lycopersicon esculentum"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="CLEG60N24"  
/clone\_lib="tomato breaker fruit"  
/tissue\_type="Pericarp"  
/dev\_stage="breaker"  
/lab\_host="SOUR"  
Site\_1: EcoRI;  
Site\_2: XhoI; supplier: Boyce Thompson Institute;

sequencing: The Institute for Genomic Research. Fruit were harvested at the breaker stage (first sign of lycopene accumulation on the blossom end of fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."

```
BASE COUNT      226 a  143 c  155 g  196 t
ORIGIN

Query Match      23.6%; Score 425.6; DB 10; Length 720;
Best Local Similarity 74.4%; Pred. No. 6.3e-116;
Matches 536; Conservative 0; Mismatches 184; Indels 0; Gaps 0;

Qy 740 ctaacgcgggttggtctatttcaatggcgcgggtattggtatgtcggagatgatttac 799
Db 1 CAAGCCGGGTTAGTCTATTTCATAAACCAGATTACTTGTCTGAAGATGATTTCG 60

Qy 800 cttaccaagttcagatcactcccaatggagatttaaaaccggttggtcggttcgattttg 859
Db 61 CTTACCATGTAAGGTAAACACCCACCGCGGATCTTAAACACAGAGGGTGCATTGCTTCG 120

Qy 860 atggacaattagaatccacaatgattgcccaccggaaagtgcgaaccgggaatccgggtgaac 919
Db 121 ACGGCCAGTAAATCCACCATGATAGCTCACCCAAAGCTCGACCCAGTTTCCGGTGAGC 180

Qy 920 tcttcgcttaagctacagctcgctttcaagccttaactaaataacttcgattctcac 979
Db 181 TATTGCTCTTAGCTPACGATGTGATTCAGAGCCCATACCTCAAGTACTTCAGATTTTCAA 240

Qy 980 cggacggaactaaatcacgcagctcgagatcagcttgatcagccaacgatgatgcacg 1039
Db 241 AAATGGGAAATCAATGATGTTGAAATCCAGTTGAGACCCACACATGATGCATG 300

Qy 1040 atttcgcgattacagagaacttcgctcgctacctgcacgcagcagctgttttcaagctgc 1099
Db 301 ATTTCCGAATTACTGAGAACTTCGTCGTCATTCTCTGTATCAACAAGTCGTTTTCAAGATGT 360

Qy 1100 cggagatgacccgggtgggtctccgggttggtttacgacaagaadaaggtcgcaagattcg 1159
Db 361 CTGAAATGATCCGTGGAGGTTTACCCGGTGGTTTACGACAAAGAACAAAGTTTCCCGATTG 420

Qy 1160 ggattttagacaaatcacgcgaagattcatcgaaacattaaagtgattgatgtccagatt 1219
Db 421 GTATTCTGGATAAGTACGCGAAAGATGGGTCTGATTTGAAATGGTTGAAGTACCTGATT 480

Qy 1220 gcttctgtccatctctggaacgcttgggaagaccgagaaacagatgaagtgcgtgga 1279
Db 481 GTTCTGTGTTCCACCTCTGGAATGCTTGGGAAGAGCAGAAACAGATGAAATCGTTGTAA 540

Qy 1280 taggctctgtatgactccaccagactcaattttcaacgagctcgcagagaatctcaaga 1339
Db 541 TTGGTTCATGTATGACACCCACAGACTCTCAATTTTCAATGAAATGTGATGAAGGGCTAAAGA 600

Qy 1340 gtgctcgtctgaaatccgctgaaatcctcaaaaccggtggaatcaactccgctccgactca 1399
Db 601 GTGTTTATCCGAATTCGCTCTCAATTTTGAACAGGGAAATTCACAGAAATCCATAA 660

Qy 1400 tctccaacgaagatcaacaagtcaacctcgaagcgggattggtcaacgaacaatgctgcg 1459
Db 661 TCGAAACCCCGATGAACAAGTGGATTTAGAAAGCTGGAATGCTGAACCGAAACAAACTCTG 720

RESULT 2
BM408615
LOCUS      BM408615
DEFINITION EST582942 tomato breaker fruit Lycopersicon esculentum cDNA clone
            cLEG45023 5' end, mRNA sequence.
ACCESSION  BM408615
VERSION    BM408615.1 GI:18260245
KEYWORDS   EST.
SOURCE     tomato.
ORGANISM   Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE 1 (bases 1 to 805)  
AUTHORS Alcala,J., Vrebalov,J., White,R., Vision,T., Karamycheva,S.A., Tsai J., Bougri,O., Kirkness,E., Utterback,T., Van Aken,S., Rönning C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.  
Generation of ESTs from tomato fruit tissue, breaker stage (2002) Unpublished (2002)

TITLE JOURNAL  
COMMENT  
Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>  
This clone is available through the Clemson University Genomics Institute

Seq primer: T3.

FEATURES  
source  
1. .805  
/organism="Lycopersicon esculentum"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="cLEG45023"  
/clone\_lib="tomato breaker fruit"  
/tissue\_type="Pericarp"  
/dev\_stage="breaker"  
/lab\_host="SOLR"  
/notes="Vector: pBluescriptSKmCuaadapt; Site\_1: EcoRI; Site\_2: XhoI; supplier: Boyce Thompson Institute; sequencing: The Institute for Genomic Research. Fruit were harvested at the breaker stage (first sign of lycopene accumulation on the blossom end of fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."  
BASE COUNT 262 a 144 c 180 g 219 t  
ORIGIN

Query Match 23.2%; Score 417; DB 10; Length 805;  
Best Local Similarity 71.8%; Pred. No. 2.5e-113;  
Matches 575; Conservative 0; Mismatches 220; Indels 6; Gaps 2;

```
Qy 899 tcgaccgcgaatccgggaactcttcgctttaagctacgcagctcggttccaagccttacc 958
Db 6 TCGACCAGTTTCCGGTGAGCTATTGCTCTTAGCTACGATGTGATTTCAGAAAGCATACC 65

Qy 959 taaaatacttcgattctcaccggacggaactaaatcacgcgacgctcgagattcagcttg 1018
Db 66 TCAAGTACTTCAGATTTTCAAAAAATGGGAAAAATCAAAATGATGTTGAAATTCAGTTG 125

Qy 1019 atcggccaacgatgatgcagcatttcgcgattacagagaacttcgctcgctgaactgacc 1078
Db 126 AAGACCAACAATGATGATGATTTCGCAATTACTGAGAACTTCGTCGTCATTCTGTATC 185

Qy 1079 agcaagtcggtttcaagctgccgagatgatccgcgctgggtctccggtggtttacgaca 1138
Db 186 AACAAAGTCGTTTTCAAGATGTCGAAATGATCCGTGGAGGTTTACCCGGTGGTTTACGACA 245

Qy 1139 agaacaaggtcgcaagattcgggatttttagaataatccgcgaagattcatcgaaacatta 1198
Db 246 AGAACAAAGTTTCCGATTTTGGTATTCTGGATAAGTAGCCGAAAGATGGGTCTGATTGA 305

Qy 1199 agtggattgatgtccagattgcttcgcttcctccatctctggaacgcttgggaagagccag 1258
Db 306 AATGGTGAAGTAGTACCTGATTGTTTCTGTTTCCACCTCTGGAATGCTTGGGAAGAAGCAG 365

Qy 1259 aaacagatgaagtgcgtgcgtagagggccctgatgactccaccagactcaattttcaacg 1318
Db 366 AACAGATGAAATCGTTGTAATGTTTCATGATATGACACCAACCAAGACTCCATTTTCAATG 425

Qy 1319 agtctgacgagaatctcaagagtgctcctgtctgaaatccgctgaatctcaaaacccgtg 1378
Db 426 AATGTGATGAAGGGCTAAAGAGTGTGTTTTATCCGAAATCCGTCCTCAATTTGAAACAGGGA 485
```

```

Qy 1379 aatcaactgcgtccgagatcctcaacgaagatcaacaagtcaacctcgaaagcgga 1438
Db 486 AATCAACAAGAAATCCATAATCGAAACCCGGATGAACAAGTCAATTTAGAAAGCTGGAA 545
Qy 1439 tggtaacagaaacatgcctcgccggtgaaaccaaattcgcttacttggcttagccgagc 1498
Db 546 TGGTAACCCGAAACAACACTCGGAAGGAAACAGAGTATGCTTATTTGGCTATCGTGAAC 605
Qy 1499 cgtggcctaaagtctcaggattcgctaaagtgtgatctcactactggagaagttaagaaac 1558
Db 606 CATGGCCAAAGTTTCTGGTTTTCGAAAGTAAACCTGTTCAACGGTGAAGTTGAGAAAT 665
Qy 1559 atctttacggcgataacggttacggaggagagcctctgtttctcccgagagaagg99ga- 1617
Db 666 TCATTATTTGTTGACAAACAATATGTTGGGGAACCTCTTTTTCACAGAGAACCCCAACAG 725
Qy 1618 ----gaggaagacgaaggtacatcctctgtcttcgttcacgacgagaacatcgaaatc 1673
Db 726 CAGGAAGAGACGATGGTTATATTTTAGCTTTTCGTTACAGATGAGAA-AGATGGGAATC 784
Qy 1674 ggagttacagatagttaaagc 1694
Db 785 AAAACTGCAAAATGGTAACGC 805

RESULT 3
LOCUS BH549344 781 bp DNA linear GSS 14-DEC-2001
DEFINITION BOGRQ53TF BOGR Brassica oleracea genomic clone BOGRQ53, DNA
sequence.
ACCESSION BH549344
VERSION BH549344.1 GI:17801124
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 781)
Town, C.D., Van Aken, S., Utterback, T. and Fraser, C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)
Other GSSs: BOGRQ53TR
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TF
Class: sheared ends.
FEATURES             Location/Qualifiers
     source            1..781
     /organism="Brassica oleracea"
     /strain="T0100DH3"
     /db_xref="taxon:3712"
     /clone="BOGRQ53"
     /clone_lib="BOGR"
     /note="Vector: pHS01; Site 1: BstXI; 2-3 kb sheared
     genomic DNA inserted into pHS01 using BstXI linkers"
BASE COUNT 204 a 203 c 192 g 182 t
ORIGIN
Query Match 23.1%; Score 416.2; DB 12; Length 781;
Best Local Similarity 70.8%; Pred. No. 4.3e-113;
Matches 553; Conservative 0; Mismatches 228; Indels 0; Gaps 0;
Qy 274 gggggggggcagcttgagacggcgagggttctcttgcagcagcagagagtaac 333
Db 1 GCAGCGCGGATTCGATCGATGGCTGAGCGAGCGTTGATCTCAGTTCAGTAACG 780

```

```

Qy 334 cggcttcctaaacggcgtgatcctcagatcgccggaaattttgctccgggtgaat 393
Db 61 CCTCTTCCCAAAACCGCTGATCCAGTATTCAAATCGCCGGGAATTTATTTCCCGGTGCG 120
Qy 394 gaacggccgctccggcgtaattcttcgggtggtcggaataacttcccgatcccatcaagga 453
Db 121 GAAATTTCCCGTCCGAGAGTGTCTCGACGTTGAAGGAACAATCCCTGACTGCAATTAACGGA 180
Qy 454 gfgtatgfcgcaacggagctaacccacttcacgagcgggtgacaggtcaccaactcttc 513
Db 181 GTTATGTCCTGTAACGGTCAAAATCCGATGTTTCGAGCCCAATCGTGGGCATCATTTGTTTC 240
Qy 514 gacggagacgggtatggttcacggcgtcgaattcgaacacgggttcagctagctacgcttgc 573
Db 241 GACGGAGACGGAATGGTTCACGCAGTTAAAAATAACCAACGGTTCAGCAAGCTACGCATGC 300
Qy 574 cggtttactcagactaacccggttttctcaggaaactcaattgggtcgacgggttttccccc 633
Db 301 CGGTTTACTTAAACCGAGAGATTAATTCAGAAAAACGATTGGGTTCGACCGGTTTTTCGG 360
Qy 634 aaagccatcggtgagcttcacggccacacccggtattgccgactcatctattctacgcc 693
Db 361 AAAGCAATCGCGAGCTTCACGCCCACTCGGGAATCGCACGTCTGATGCTTTTTTACGGA 420
Qy 694 agagctgcagccggtatagtcgacccgggcacacggacccggtgttagctaacgccggtttg 753
Db 421 CGTGGGCTTTGCGGCTCGTCAACAATCAAAACGGCGTCGGGGTAGCGAACCCTCGGTTTG 480
Qy 754 gtcatttcaatggcgggttattggtctatgtcggaggatgatttaccttaccgaagtcag 813
Db 481 GTTACTTCAATAACCGGCTTTTAGCAATGTGAGAAGACGATTTACCGTACCAATTAATAA 540
Qy 814 atcactcccaatggagatttaaaaaccgttggtcgggtcgattttgatggacaattagaa 873
Db 541 ATAACTCAAAACCGGACCTCCAAACCGTTGGCGTTACGATTCGACGGTCAGTTAAAA 600
Qy 874 tcacaatgattgcccaccccgaaagtgcacccgggaatccgggtgaactctctcgtttaagc 933
Db 601 TCATCGATGATCGCCACCCCGAAACTCGACCCGGTTACAAAGGAGCTCCACCGCTAAGC 660
Qy 934 tacgagctcggttcaaaccttacctaaataacttccgatttcccgagacggaactaaa 993
Db 661 TACGACGTCGTTAAAGAGCTTACCTGAAATATTTTTCAGATTCGCGCGACGGCGTTAA 720
Qy 994 tcacggagcgtcgagattcagcttgatcagcaacgagatgacgagatttcgagattaca 1053
Db 721 TCACCGAGCTGGAGATCCCGCTCGAGACTCCGACGAGTTCAGGATTCGCTATAACG 780
Qy 1054 g 1054
Db 781 G 781

```

```

RESULT 4
LOCUS AW933524
DEFINITION ESN35283 tomato fruit mature green, TAMU Lycopersicon esculentum
CDNA clone cLEP54E2 5', mRNA sequence.
ACCESSION AW933524
VERSION AW933524.2 GI:11388127
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 643)
REFERENCE 1
AUTHORS Alcala, J., Vrebalov, J., White, R., Matern, A.L., Holt, I.E., Liang, F.,
Upson, J., Hansen, T., Craven, M.B., Bowman, C.L., Ahn, S., Rongai, C.M.,
Fraser, C.M., Martin, G.B., Tanksley, S.D. and Giovannoni, J.
Generation of ESTs from tomato fruit tissue
TITLE

```



```
JOURNAL Unpublished (1999)
COMMENT On May 30, 2000 this sequence version replaced gi:8108841.
Contact: CUGI
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.

FEATURES
    source Location/Qualifiers
        1..643
            /organism="Lycopersicon esculentum"
            /cultivar="TA496"
            /db_xref="taxon:4081"
            /clone="cLEF54E2"
            /clone_lib="tomato fruit mature green, TAMU"
            /tissue_type="fruit pericarp"
            /dev_stage="mature green (3-5 days pre-ripening)"
            /lab_host="SOLR"
            /note="vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
            XhoI; cLEF - Fruit were tagged at the 1cm stage and
            harvested 3-5 days prior to ripening. Fruit were cut in
            half to verify the seeds were indeed 'immature' and the
            seeds and locules were discarded prior to freezing the
            pericarp"
BASE COUNT 184 a 130 c 142 g 186 t 1 others
ORIGIN

Query Match 21.3%; Score 382.8; DB 9; Length 643;
Best Local Similarity 74.7%; Pred. No. 3.8e-103;
Matches 480; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

Qy 685 ttctacgcagagctgcagccggtatagtcgaccggcacagcggaaccgggtgtagctaac 744
Db 1 TTTNACGTCGTGGGCTCTTCGGACTTGTGATCACAGTAAGGAACCTGGTGTGCAAC 60

Qy 745 gccggtttggctctattcaatggccggtattggtatgtcggaggatgattaccttac 804
Db 61 GCCGGTTTAGTCTATTTCAATAACCGATTACTTGTGTATGTCGAGATGATTGCGTTAC 120

Qy 805 caagttcagatcactcccaatggagatttaaaacccgttgctggttcgatttgatgga 864
Db 121 CATGTAAAGGTAACACCCACCGCGCATCTTAAACAGAGGGTCGATTGATTCGATTCGACGGC 180

Qy 865 caattagatcaccaatgatgcaccaccggaagtgcaccgggaatccggggaactcttc 924
Db 181 CAGCTTAAATCCACCATGATAGCTCACCCAAAGCTCGACCCAGTTTCGGGTGAGCTATT 240

Qy 925 gctttaagctacgagctgcttcaagccttacctaaataactcttcgattctcacggac 984
Db 241 GCTCTTAGCTACGATGTGATTCAGAAGGCATACCTCAAGTACTTCAGATTTTCAAAAAT 300

Qy 985 ggaactaaatcacccgagctcgagattcagcttgatcgagcccaacgatgatgcagatttc 1044
Db 301 GGGGAAAATCAATGATGTTGAAATTCAGTTGAAGACCCCAACAAATGATGATGATTC 360

Qy 1045 gcgattacaggaacttcgctcgtcgtactgcaccgcaagtcggtttcaagctgcggag 1104
Db 361 GCAATTACTGAACTTCGTCGTCTCATCTCGATCAACAAGTCGTTTCAAGATGCTGAA 420

Qy 1105 atgattcccggtgggtctccggtggtttacgacaagaacaagtcgcaagattcgggatt 1164
Db 421 ATGATCCGTGGAGGTTACCCGTTGTTTACGACAAGAACAAAGTTTCCGATTTGGTATT 480

Qy 1165 ttagacaataccgcaagattcatcgaaacttaagtggtgattgatgtccagattgcttc 1224
Db 481 CTGGATAAGTACGCCAAAGATGGGTCTGATTTGAAATGGGTTGAAGTACCTGATTGTTTC 540

Qy 1225 tgcctccatctctggaacgcttggaagagccagaaaaacagatgaagtcgctgtagagg 1284
Db 541 TGTTTTCCACTCTGGAATGCTTTGGGAAGAAGCAGAAAACAGATGAAATCGTTGTAAATGGT 600

Qy 1285 tcctgtatgaactcccagagactcaatttttcaacgagctctgacg 1327

Db 601 TCATGTATGACACCACCAGACTCCATTTTCAATGAATGTGATG 643

RESULT 5
BM408565 787 bp mRNA linear EST 22-JAN-2002
LOCUS EST582892 tomato breaker fruit Lycopersicon esculentum cDNA clone
DEFINITION cLEG45E23 5' end, mRNA sequence.
ACCESSION BM408565
VERSION BM408565.1 GI:18260195
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 787)
AUTHORS Alcala,J., Vrebalov,J., White,R., Vision,T., KaramyCheva,S.A., Tsai
,J., Bougri,O., Kirkness,E., Utterback,T., Van Aken,S., Roming
,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
Generation of ESTs from tomato fruit tissue, breaker stage (2002)
Unpublished (2002)
TITLE Generation of ESTs from tomato fruit tissue, breaker stage (2002)
JOURNAL
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
This clone is available through the Clemson University Genomics
Institute
Seq primer: T3.
FEATURES
    source Location/Qualifiers
        1..787
            /organism="Lycopersicon esculentum"
            /cultivar="TA496"
            /db_xref="taxon:4081"
            /clone="cLEG45E23"
            /clone_lib="tomato breaker fruit"
            /tissue_type="pericarp"
            /dev_stage="breaker"
            /lab_host="SOLR"
            /note="Vector: pBluescriptSKmCuaadapt; Site 1: EcoRI;
            Site 2: XhoI; supplier: Boyce Thompson Institute;
            sequencing: The Institute for Genomic Research. Fruit
            were harvested at the breaker stage (first sign of
            lycopene accumulation on the blossom end of fruit). Fruit
            were cut in half and the seeds and locules were discarded
            prior to freezing the pericarp."
BASE COUNT 216 a 194 c 165 g 212 t
ORIGIN

Query Match 20.9%; Score 375.6; DB 10; Length 787;
Best Local Similarity 67.8%; Pred. No. 6.2e-101;
Matches 525; Conservative 0; Mismatches 249; Indels 0; Gaps 0;

Qy 205 tctcccgcattgttgttaagcccaagcccaagaatcccaactaaacagatgaatttg 264
Db 5 TCACACCCAAAACAAGAAAAACAACAACCTCTCTTCTTCAACTTCCACAGTGGAAATTA 64

Qy 265 ttcagagagcggcgccgagcgttggagcggcgagggtttccctgtcagccagag 324
Db 65 GTGCAGAAAGCAGCAGCAATGCGCTTTAGATGCTGTAGAAAGTGCTTTAACTTAACATGAA 124

Qy 325 aagctacacccgctctctaaacgggtgatcctagtggttcagatcgccggaatttgct 384
Db 125 CTTGAACACCCCTTTTGGCGAAAAACAGCCGACCCACAGGTCACAGATTTCTGGGAAATTTTCT 184

Qy 385 ccggtgaatgaacggcgcgtccggcgtaattctccggtggtcggaataactcccgattcc 444
Db 185 CCGGTACCGGAAAAATCCAGTCTGTCAATCTCTTCGGTTCACCGGAAAAAATACCCAAATGT 244
```

QY 445 atcaaaaggagtgtatgtgcgaacggagagtaaacacacttcacagagcgggtgacaggtaac 504  
 Db 245 GTTCAAGGGGTTTACGTTTCAAAACGAGAGTAACCTCTTTTGAACCAACCCGACAC 304  
 QY 505 caatttttcagcggagacggatggttcacgcgcgtcaaatcgaacacgggttcagctagc 564  
 Db 305 CATTCTTCGACGGGACGGTATGTTTCACGCGTTCATTTCAAAATGGTTCGGCTAGT 364  
 QY 565 tacggttcggtttactcagactaacacgggtttgttcaggaacgtcaattgggtcgaccg 624  
 Db 365 TAGCCTTGCGGTTTCACTGAACAGAGAGGCTTGTTCAAGAAAGCTTTGGTTCGCCCT 424  
 QY 625 gttttcccaagcattcggtagcttcacgcccacacgggtatgcccagatcattgcta 684  
 Db 425 GTTTTCCCTTAAAGCCATTGGTGAATTCAGTGTCTACTCTGGAATTCGAAGGCTTATGCTG 484  
 QY 685 ttctacgcagagcgcagccggtatagtcgaccgggcacacggaacgggtgtagctaac 744  
 Db 485 TTTTACGCTCGTGGGCTCTTCGACTTGTTGATCACAGTAAGGAAGCTGGTGTTCGAAC 544  
 QY 745 gccggtttggtctatttcaatggcgggttattggctatgctcggaggtgatttaccctac 804  
 Db 545 GCCGGTTTGTCTATTTCAATACCGATTACTTGTCTGAGAGATGATTTGCCCTTAC 604  
 QY 805 caagttcagatcactcccaatggagatttaaaacggttggtcggttcgattttgatgga 864  
 Db 605 CATGTAAAGGTAAACACCCCGCGATCTTAAACACAGAGGGTCGATTTCGATTCGACGCG 664  
 QY 865 caattagatcaccaatgattccacccgaaagtcgaccggaatccggtgaactcttc 924  
 Db 665 CAGCTAAATCCACCATGATAGTCTACCCAAAGCTCGACCCGATTTCCCGGTGAGCTATT 724  
 QY 925 gtttaagctacgagctgtttcaagccttaccataaatacttccgattctca 978  
 Db 725 GCTCTTAGCTACGATGTGATTTCAGAGGCCATACCTCCGACTTTCAGATTTTCA 778

RESULT 6  
 LOCUS BF113346 592 bp mRNA linear EST 18-MAY-2001  
 DEFINITION EST440936 tomato breaker fruit Lycopersicon esculentum cDNA clone  
 CLEG44B9 5' sequence, mRNA sequence.  
 ACCESSION BF113346  
 VERSION BF113346.1 GI:10943036  
 KEYWORDS EST.  
 SOURCE tomato.  
 ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;  
 Lycopersicon.  
 REFERENCE 1 (bases 1 to 592)  
 AUTHORS Alcalá, J., Vrebalov, J., White, R., van der Hoeven, R.S., Holt, I.E.,  
 Liang, F., Hansen, T.S., Craven, M.B., Bowman, C.L., Ronning, C.M.,  
 Nierman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley,  
 S.D.  
 TITLE Generation of ESTs from tomato fruit tissue, breaker stage  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: CUGI  
 Clemson University  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>.  
 Location/Qualifiers  
 FEATURES  
 source  
 1..592  
 /organism="Lycopersicon esculentum"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="CLEG44B9"  
 /clone\_lib="tomato breaker fruit"  
 /tissue\_type="pericarp"  
 /dev\_stage="breaker"  
 /lab\_host="SOLR"

/note="Vector: pBluescriptSKmCUDapt; Site 1: EcoRI;  
 Site 2: XhoI; supplier: Boyce Thompson Institute;  
 sequencing: The Institute for Genomic Research. Fruit  
 were harvested at the breaker stage (first sign of  
 lycopene accumulation on the blossom end of fruit). Fruit  
 were cut in half and the seeds and locules were discarded  
 prior to freezing the pericarp."

BASE COUNT 194 a 106 c 131 g 161 t

Query Match 19.1%; Score 344; DB 10; Length 592;  
 Best Local Similarity 73.8%; Pred. No. 1.5e-91;  
 Matches 437; Conservative 0; Mismatches 155; Indels 0; Gaps 0;  
 QY 893 cgaagtcgacccggaatcccggtgaactcttcgctttaaagctacgacgtcggttcaaacg 952  
 Db 1 CAAAGCTCGACCCAGTTTCCGGTGAGCTATTGCTCTTAGCTACGATGTGATTGAGAAGC 60  
 QY 953 cttacctaaaatacttccgattctccacggacgggaactaaatcacccgacgtcgagattc 1012  
 Db 61 CATACCTCAAGTACTTTCAGATTTTCAAAAAATGGGAAAAATCAAATGATGTTGAAATTC 120  
 QY 1013 agcttgatcagccaacgatgatgcacgatttcgagatttcagagaaacttcgctcgctac 1072  
 Db 121 CAGTTGAAGACCCAAATGATGATGATTTTCGCAATTTACTGAGAACTTCGTCGTCATT 180  
 QY 1073 ctgaccagcaagtcgttttcaagctgccgagatgacgcggtgggtctccggtgggttt 1132  
 Db 181 CTGATCAACAAGTCGTTTTCAAGATGCTCGAAATGATCCGTGGAGGTTACCCGGTGGTT 240  
 QY 1133 acgacaagaacaagtcgcaagattcgggattttagacaaataacccggaagattcatcga 1192  
 Db 241 ACGACAAGAACAAAGTTTCCCGATTGTTGATTTCTGGATAAGTACCGGAAAGATGGTCTG 300  
 QY 1193 acatttaagtgattgattgctccagatttctctcgttcctccatctctggaaacgttgggaag 1252  
 Db 301 ATTTGAAATGGGTTGAAGTACCTGATTTCTGTTTCCACCTCTGGAATGCTTGGGAAG 360  
 QY 1253 agcagaacaagatgaagtcgctgagatagggtcctgactccaccagactcaattt 1312  
 Db 361 AAGCAGAAACAGATGAATCGTTGTAATGGTTCATGATGACACACAGACTTCATTT 420  
 QY 1313 tcaacgagctctgacgagaatctcaagagtcctgtcgtgaatccgctgaaatcctcaaaa 1372  
 Db 421 TCAATGAATGTGATGAAGGCTTAAAGAGTGTGTTTATCCGAAATCGTCTCAATTTGAAA 480  
 QY 1373 ccggtgaatcaactccgctccgattcctccaaagagatcaacaagtcacacctcgaag 1432  
 Db 481 CAGGGAATTCACAAAGAAATCCATAATCGAAACCCCGATGACCAAGTGAATTTAGAAG 540  
 QY 1433 caggatggtcaacagaacatgctcggcgttaaaaccaaactcgttactt 1484  
 Db 541 CTGGAATGGTGAACCGAAACAAACTCGGAAGAAACAGAGTATGCTTATTT 592

RESULT 7  
 LOCUS BE461924 566 bp mRNA linear EST 18-MAY-2001  
 DEFINITION EST413439 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA  
 clone CLEG40L2, mRNA sequence.  
 ACCESSION BE461924  
 VERSION BE461924.1 GI:9506322  
 KEYWORDS EST.  
 SOURCE tomato.  
 ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;  
 Lycopersicon.  
 REFERENCE 1 (bases 1 to 566)  
 AUTHORS Alcalá, J., Vrebalov, J., White, R., van der Hoeven, R.S., Holt, I.E.,  
 Liang, F., Hansen, T.S., Craven, M.B., Bowman, C.L., Ronning, C.M.,  
 Nierman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley,  
 S.D.

Nierman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.

TITLE Generation of ESTs from tomato fruit tissue, breaker stage  
JOURNAL Unpublished (2000)  
COMMENT Contact: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

## FEATURES

Location/Qualifiers

1. 566

/organism="Lycopersicon esculentum"

/cultivar="TA496"

/db\_xref="taxon:4081"

/clone="cUEG40L2"

/clone.lib="tomato breaker fruit, TIGR"

/tissue type="pericarp"

/dev stage="breaker"

/lab\_hosts="SOLR"

/notes="Vector: pBluescriptSKMvUadapt; Site 1: EcoRI; Site 2: XhoI; Fruit were harvested at the breaker stage (first sign of lycopenic accumulation on the blossom end of the fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."

184 a 104 C 124 G 154 T

BASE COUNT

ORIGIN

Query Match 18.5%; Score 332.4; DB 10; Length 566;

Best Local Similarity 74.2%; Pred. No. 4.3e-88;

Matches 420; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

Qy 893 cgaagtcgacccggaatcccggtgaactcttcgctttaagtaagacgtcgtttcaagc 952

Db 1 CAAGCTCGACCAGTTCGCGTGAGCTATTGCTCTTAGCTACGATGTGATTCAGAAGC 60

Qy 953 cttaactaaatacttcgattctcaccggacgggaactaaatcaccggacgtcgagattc 1012

Db 61 CATACCTCAAGTACTTCAGATTTTCAAAAAATGGGAAAAATCAATGATGTGAAATTC 120

Qy 1013 agcttgatcagcaacgatgatgcacgatcttcgggatttcacagaaactctgcgtcgac 1072

Db 121 CAGTTGAAGACCCACATGATGATGATTTTCGCAATTTACTGAGNACTTCGTCGTATTC 180

Qy 1073 ctgaccagagtcgttttcaagctgccggagatgatccgcggtgggtctccggtggtt 1132

Db 181 CTGATCAACAAGTCGTTTCAAGATGTCGAAATGATCCGTCGAGGTTTCACCGTGGTGT 240

Qy 1133 acgacaagaacagtcgaagattcgggattttagacaaatagccgaagattcatoga 1192

Db 241 ACGACAAGAACAAAGTTTCCCGATTGTGATTTCTGGATAAGTACGCGAAAGATGGTCTG 300

Qy 1193 acattaaatggattgatctccagattctctctgcttccatctctggacgcttgggaag 1252

Db 301 ATTGAAATGGTTGAAGTACCTGATGTTCTGTTTCCACTCTGGAATCTGTTGGGAG 360

Qy 1253 agccgaacagatgaagtcgctgtagaggttcctgtagactccaccagactcaattt 1312

Db 361 AAGCAGAACACAGATGAATCGTTGTAATGGTTTCATGTATGACACACACAGACTCCATT 420

Qy 1313 tcaacgactcgacgagaaatctcaagatgtccctgctgaaatccgctgaaatcctcaaaa 1372

Db 421 TCAATGAATGTGATGAAGGGCTAAAGAGTGTGTTTATCGAAATCCGTTCAATTTGAAAA 480

Qy 1373 cgggtgaatcaactcgccgtccgcatctcctcaacgaagatcaacaagtcacactcgaag 1432

Db 481 CAGGGAATCAACAGAAAAATTCATAATCGAAAAACCCGGATGACAAAGTATTTAGAG 540

Qy 1433 cagggatgggtcaacagaacatgcttc 1458

Db 541 CTGGAATGCTGAACCGAAACAAACTC 566

## RESULT 8

BM085672

LOCUS

DEFINITION

BM085672 553 bp mRNA linear EST 19-NOV-2001  
saJ28a02.y1 Gm-cl066 Glycine max cDNA clone SOYBEAN CLONE ID:  
Gm-cl066-3867 5', similar to TR:O24023 O24023 NEOXANTHIN CLEAVAGE  
ENZYME. ; mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Glycine max

soybean.

1 (bases 1 to 553)  
Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V., Khanna  
A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,  
Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers  
Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk  
R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann  
R., Waterston, R. and Wilson, R.  
Public Soybean EST Project  
Unpublished (1999)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: [est@watson.wustl.edu](mailto:est@watson.wustl.edu)

This clone is available through: ResGen, Invitrogen Corp. 2130

South Memorial Parkway Huntsville, AL 35801 For further information

call: (800)-533-4363 or contact: [ccu@resgen.com](mailto:ccu@resgen.com) web site:

[www.resgen.com](http://www.resgen.com)

High quality sequence stop: 426.

Location/Qualifiers

1. 553

/organism="Glycine max"

/db\_xref="taxon:3847"

/clone="SOYBEAN CLONE ID: Gm-cl066-3867"

/clone.lib="Gm-cl066"

/tissue type="Leaf and shoot tip, salt stressed, 2 week

old seedling"

/lab\_host="DH10B"

/notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:

XhoI; The cDNA library was constructed from mRNA isolated

from unexpanded leaves and the shoot tips of 2 week old

seedling from the cultivar Williams. The 2 week old

seedlings were salt stressed in a solution of 500mM NaCl

for 3 days prior to harvesting. Complementary DNA was

synthesized from mRNA using a primer consisting of a

poly(dT) sequence with a XhoI restriction site. EcoRI

adapters were ligated to the blunt-ended cDNA fragments

followed by XhoI digestion. The cDNA fragments were

directionally cloned into the EcoRI-XhoI restriction site

of the pBluescript vector. The ligated cDNA fragments were

transformed into DH10B host cells (GibcoBRL). This library

was constructed in the laboratory of Dr. Randy

Shoemaker."

148 a 165 C 120 G 120 T

BASE COUNT

ORIGIN

Query Match 18.0%; Score 323.2; DB 10; Length 553;

Best Local Similarity 74.1%; Pred. No. 2.4e-85;

Matches 409; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

Qy 822 caatggagatttaaaaccgttgctcggttcgattttgatggacaattagaaatcacaat 881

Db 1 CAACGGCGACTTAACCCACCGTCGCGCTTACAACTTTTAACGGCCAGTTAAAAATCACAAT 60

```
QY 882 gattgccaccggaagtcgaccggaatccggtgaactcttcgctttaagctacgaagt 941
Db 61 GATCGCTACCCCAAACTGACCCCGTGAGCAACAACCTCCACGCCCTCTCTACGACGT 120
QY 942 cgtttcaagccttacctaaataacttcgattctcaccggagcggaactaaatcacgga 1001
Db 121 CGTCCAGAGCCATATCTCAAGTACTTCGGCTTCTCCCGAGGGGGTCAATCCCCCGA 180
QY 1002 cgtcgagattcagcttgatcagccaacgatgatgcagatttcgctgattacagagaactt 1061
Db 181 CGTCGAAATTCCTTGAAGAAACCCACCATGATGCACGATTCGCCCATCACCGAGAACTT 240
QY 1062 cgtcgctacactgaccacaagtcgttttcaagctgcggagatgatccgggtgggtc 1121
Db 241 CGTCGTGTCGCCCGACGACGAGGTGCTCTCAAGTTATCAGAAATGATCACCGGAGGCTC 300
QY 1122 tccgggtggttacgacaagaacaggctcgcaagattcgggattttagacaaatacgcga 1181
Db 301 CCCCGTTGTACGACAGACAGAGTTTCCAGATTCGGGATTCGCAAGAAGATGCTTAA 360
QY 1182 agattcatcgaaacattaaatgattgatctccagattgcttcttcttccatctctggaa 1241
Db 361 AGATGCAAAATGATGAAATGGATCATGCTCCGAGTGTCTTCTGTTCCACTCTGGAA 420
QY 1242 cgttgggaagacgagaacagatgaagtctcgtgatagggtcctgtatgactccacc 1301
Db 421 CGCTGGGAGGAGCGGAAATGATGAGATCGTTGTTCATCGGTTCCTGATGACCCCGC 480
QY 1302 agactcaatttcaagagctcgacagagaatctcaagagtgctcgtctgaaatccgct 1361
Db 481 GCATCCATTTTCAAGATTCGAGAGAGTGTGAGAGCACTTGTGCGAGATTAAGGT 540
QY 1362 gaattcctcaaac 1373
Db 541 GAATTTGAAGAC 552

RESULT 9
BM085005 saj31a08.y1 Gm-c1066 Glycine max cDNA clone SOYBEAN CLONE ID:
LOCUS Gm-c1066-4240 5' similar to TR:024023 024023 NEOXANTHIN CLEAVAGE
DEFINITION ENZYME.; mRNA sequence.
ACCESSION BM085005 linear EST 19-NOV-2001
VERSION BM085005.1 GI:16995633
KEYWORDS soybean.
SOURCE Glycine max
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE 1 (bases 1 to 559)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna
,A., Bolla,B., Marr,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact: ccu@resgen.com web site:
www.resgen.com
High quality sequence stop: 421.
```

FEATURES  
source

Location/Qualifiers  
1..559  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="SOYBEAN CLONE ID: Gm-c1066-4240"  
/clone\_lib="Gm-c1066"  
/tissue\_type="Leaf and shoot tip, salt stressed, 2 week old seedling"  
/lab\_host="DH10B"  
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from mRNA isolated from unexpanded leaves and the shoot tips of 2 week old seedling from the cultivar Williams. The 2 week old seedlings were salt stressed in a solution of 500mM NaCl for 3 days prior to harvesting. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

BASE COUNT 132 a 206 c 105 g 116 t  
ORIGIN

Query Match 17.7%; Score 118; DB 10; Length 559;  
Best Local Similarity 73.1%; Pred. No. 8.7e-84;  
Matches 408; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 628 ttcccaaaagccatcggtgagcttcacggccacacccggtattgcccgactcatgtattc 687  
Db 2 TTCCCAAGCCATCGGTGAATCCACGCCACTCCGCGATCGCCCGCTCTCTCTTC 61  
QY 688 tacgccagagtgcagccggtatagtcgaccggcacacggaacccggtgtagtaacgcc 747  
Db 62 TACGCCCGCTCCTCTTCGGGCTCGTCGATGGTCCACGGCATGGGTGTGCCAACGCC 121  
QY 748 ggttggtctattcaatggcggttattgctatgtcggaggatgattacccttaccaa 807  
Db 122 GCGCTCGTCTACTTCAACAACCATCTCTTAGCCATGTCCGAAGATGACTTACCCTACCAT 181  
QY 808 gtccagatcactcccaatggagatttaaaacccgttgctcggttcgattttgatggacaa 867  
Db 182 CTCCGCATAACACCAACCGCGACTTACCACCGTCGCGCGTTACAACCTTTAACCGCCAG 241  
QY 868 ttagaatccacaatgattgcccaaccgaaagtgcaccgggaatccggtgaactctcgct 927  
Db 242 TTAATAATCCAAATGATCGCTCACCCCAAACTCGACCCCGTGACCAACAACCTCCACGCC 301  
QY 928 ttaagctacgagctggtttcaagcctttacataaattccgattctcccggaacgga 987  
Db 302 CTCTCTACGACGTGCTCCAGAACCATATCTCAAGTACTTCCGGTCTTCCCGACAGCGC 361  
QY 988 actaaatcacggagcgtcgagattcagttgatcagccaacgatgatcagcagatttcgcg 1047  
Db 362 GTCAAATCCCCGAGCTCGAAATTCCTTTGAAGAACCCACCATGATGATGATTCGCC 421  
QY 1048 attacagagaacttcgctcgctacactgaccagcaagtctgtttcaagctccggagatg 1107  
Db 422 ATACCGAGAACTTCGTGCTCGTCCCGACCGAGGTGGTCTTCAAGTTATACAAATG 481  
QY 1108 atccggtggttccggtggttttagacaagaacaaaggtcgcaagatttcgggatttta 1167  
Db 482 ATACCGGAGGCTCCCCCGTTGCTTACGACAAGAACAGGTTTCCAGATTCCGGATTCTC 541  
QY 1168 gacaaatcacggcaagt 1185  
Db 542 GACAAGAAATGCTAAAGAT 559

RESULT 10  
 BE432853  
 LOCUS  
 DEFINITION EST39478 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA  
 clone cLEG11G7, mRNA sequence.  
 ACCESSION  
 VERSION BE432853  
 KEYWORDS  
 SOURCE EST.  
 ORGANISM  
 Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;  
 Lycopersicon.  
 1 (bases 1 to 520)  
 Alcalá, J., Vrebalov, J., White, R., van der Hoeven, R.S., Holt, I.E.,  
 Liang, F., Hansen, T.S., Craven, M.B., Bowman, C.L., Ronning, C.M.,  
 Nierman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley  
 S.D.  
 TITLE Generation of ESTs from tomato fruit tissue, breaker stage  
 JOURNAL  
 COMMENT Unpublished (2000)  
 Contact: CUGI  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 5 prime sequence.  
 FEATURES  
 source  
 1..520  
 /organism="Lycopersicon esculentum"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="cLEG11G7"  
 /clone\_lib="tomato breaker fruit, TIGR"  
 /tissue\_type="pericarp"  
 /dev\_stage="breaker"  
 /lab\_host="SOLR"  
 /note="Vector: pBluescriptSKmCudapt; Site 1: EcoRI;  
 Site 2: XhoI; Fruit were harvested at the breaker stage  
 (first sign of lycopene accumulation on the blossom end  
 of the fruit). Fruit were cut in half and the seeds and  
 locules were discarded prior to freezing the pericarp."  
 BASE COUNT 156 a 100 c 116 g 148 t  
 ORIGIN  
 Query Match 17.6%; Score 316.6; DB 10; Length 520;  
 Best Local Similarity 75.9%; Pred. No. 2.2e-83;  
 Matches 391; Conservative 0; Mismatches 124; Indels 0; Gaps 0;  
 Qy 829 gatttaaaacccgttgctggttcatttggatggacaaattagatccacaaatgattgcc 888  
 Db 2 GATCTTAAACAGAGGGTGGATTCGATTTCGACGGCCAGCTAAATCCACCATGATAGCT 61  
 Qy 889 caccgaaagtcgaccgaatccggtgaactcttcgctttaaagctacgagtcgtttca 948  
 Db 62 CACCCAAAGCTCGACCCAGTTTCCGGTGAGCTATTTCCTTTAGTACGATGTAATCAG 121  
 Qy 949 aagccttaactaaaatacttcgattctccacggcgaggaactaaatcacggagctcgag 1008  
 Db 122 AAGCCATACCTCAAGTACTTCAGATTTTCAMAAATGGGAAATCAATGATGTTGAA 181  
 Qy 1009 attcagcttgatcagccaacgatgatgacgatttcgcgattacagagaacttcctgc 1068  
 Db 182 ATTCCAGTTGAAGACCCCAACAATGATGATGATTTTCGCAATTACTGAGAACTTCGTCGTC 241  
 Qy 1069 gtacctgaccagcaagtcgttttcagctccgagagatgacccggtgggtctccggtg 1128  
 Db 242 ATTCTGTATCAACAAGTCGTTTCAAGATGCTGAAATGATCCGGTGGAGGTTTCAACCGGTG 301  
 Qy 1129 gtttacgacaagaacaaagtcgcaagattcgggatttttagacaaatacgcggaagtcca 1188  
 Db 302 GTTTACGACAAGAAACAAAGTTTCCCGATTGTTGTTATTTCTGGATAAGTACGCGAAAGATGGG 361

Qy 1189 tcgaacattaaagtggattgatgtccagattgcttctgcttccatctcttgaaagcgttg 1248  
 Db 362 TCTGATTGAAATGGTTGAAGTACCTGATGTTTCTGTTTCCACCTCTGGAATGCTTGG 421  
 Qy 1249 gaaagccagaacagatgaagtctgctgtagaggttcctgtagtaccaccagactca 1308  
 Db 422 GAAGAAGCAGAAACAGATGAATCGTTGTAATTGTTTCATGTATGACACACCAGACTCC 481  
 Qy 1309 attttcaacagctgcagagaatctcaagagtgt 1343  
 Db 482 ATTTTCAATGAATGTGATGAAGGCGCTAAAGAGTGT 516  
 RESULT 11  
 BH559796  
 LOCUS  
 DEFINITION BOHR47TF BOHR Brassica oleracea genomic clone BOHR47, DNA  
 sequence.  
 ACCESSION  
 VERSION BH559796  
 KEYWORDS  
 SOURCE GSS.  
 ORGANISM  
 Brassica oleracea.  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.  
 1 (bases 1 to 742)  
 Town, C.D., Van Aken, S., Utterback, T. and Fraser, C.M.  
 TITLE Whole genome shotgun sequencing of Brassica oleracea  
 JOURNAL Unpublished (2001)  
 COMMENT Other GSSs: BOHR47TR  
 Contact: Chris Town  
 TIGR  
 9712 Medical Center Drive, Rockville, MD 20850, USA.  
 Tel: 301-838-3523  
 Fax: 301-838-0208  
 Email: [cdtown@tigr.org](mailto:cdtown@tigr.org)  
 DNA is from a doubled haploid provided by Tom Osborn.  
 Seq primer: TP  
 Class: Sheared ends.  
 FEATURES  
 Location/Qualifiers  
 source  
 1..742  
 /organism="Brassica oleracea"  
 /strain="TOL000DH3"  
 /db\_xref="taxon:3712"  
 /clone="BOHR47"  
 /clone\_lib="BOHR"  
 /note="Vector: pHOS1; Site 1: BstXI; 2-3 kb sheared  
 genomic DNA inserted into pHOS1 using BstXI linkers"  
 BASE COUNT 190 a 235 c 117 g 200 t  
 ORIGIN  
 Query Match 17.4%; Score 313.8; DB 12; Length 742;  
 Best Local Similarity 84.1%; Pred. No. 1.9e-82;  
 Matches 354; Conservative 0; Mismatches 67; Indels 0; Gaps 0;  
 Qy 1 atggctttttcacggcaacggctgcggtttctgggagatggcttggtggcaatcatact 60  
 Db 321 ATGGCTTCTTTTCACGGGACTACGGCGTTTCTCGAGATGGTGGTAAATCATACT 380  
 Qy 61 cagcgcgcattatcgctttctcaagctccgacttgagttattgtagctccttaccctatg 120  
 Db 381 AAACCAACCATTTATCGTCTCTTCTCAAAGCTCCGCCCTTGGGTTATTCTTCGCTGCCATG 440  
 Qy 121 gccagtcgtgcacacgaagtcgaatttcattctcgcttcacactcctccagctctt 180  
 Db 441 ACAATCGTTCCTCCACGAAAGCTCAATGTTTCTTCGGCTTCACACTCATCTCTGCTCTC 500  
 Qy 181 catttccctaaagcaatcatcaaacctctccgcattctcctgtaagccaaagccaaagaa 240  
 Db 501 CATTTCCCAAGCAATCTCCACCTCTCCCGCCATTGTTGTGAACCCCAACCAAGAA 560







QY 911 ccggtgaactcttcgctttaagctacagcgtcgcttcaaaagccttaccctaaatacttc 970  
 |||||  
 Db 481 CCcGTGAGCTATTGTCTTAGCTACGATGTGATTcAGAAGCCATACCTCAAGTACTTCA 540  
 |||||  
 QY 971 gattctc 977  
 |||||  
 Db 541 GATTCTC 547  
 |||||

RESULT 15  
 BE458861  
 LOCUS EST14153 tomato developing/immature green fruit Lycopersicon  
 DEFINITION esculentum cDNA clone CLEM4F10, mRNA sequence.  
 ACCESSION BE458861  
 VERSION BE458861  
 KEYWORDS EST.  
 SOURCE BE458861.1 GI:95031163  
 ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;  
 Lycopersicon.  
 REFERENCE 1 (bases 1 to 617)  
 AUTHORS Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,  
 Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,  
 Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley  
 ,S.D.  
 TITLE Generation of ESTs from tomato fruit tissue, immature green  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: CUGI  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 5 prime sequence.

FEATURES  
 source  
 1..617  
 /organism="Lycopersicon esculentum"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="CLEM4F10"  
 /clone\_lib="tomato developing/immature green fruit"  
 /tissue\_type="fruit"  
 /dev\_stage="immature green (5-35 days post-anthesis)"  
 /lab\_host="SOLR"  
 /note="Vector: pBluescriptSKmCuaapt; Site 1: EcoRI;  
 Site 2: XhoI; Fruit were tagged at 5 dpa (0.5 cm) and  
 harvested at 7 day intervals through 35 dpa. Equal masses  
 of tissue from each stage were combined (including seeds  
 and locules) prior to mRNA isolation."  
 BASE COUNT 164 a 145 c 137 g 171 t  
 ORIGIN

Query Match 16.9%; Score 303.4; DB 10; Length 617;  
 Best Local Similarity 68.6%; Pred. No. 2.2e-79;  
 Matches 418; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

QY 247 actaacagatgaattgttccagagagcgcgccgagcgttgagcgcggcgaggggt 306  
 |||||  
 Db 9 ACTTCCAAGTGGAAATTAGTCGAGAAAGCAGCAGCAATGGCTTAGATGCTGTAGAAAGT 68  
 |||||  
 QY 307 ttccitgtcagcagagagagtaaccgcttcctaaaaacggctgacctagtgttcag 366  
 |||||  
 Db 69 GGTTTAACTAAACATGAACCTTGAACACACCTTTGCCGAAAACAGCCGACGAGTCCAG 128  
 |||||  
 QY 367 atcgccggaaatttgcctccgtggaatgaacgcccgcgcgtgaatcttcgggtggtc 426  
 |||||  
 Db 129 ATTTCTGGGAATTTTGCTCCGGTACCGGAAATCCAGTCTGTCAATCTCTTCGGTCAAC 188  
 |||||  
 QY 427 ggaaacttcccgattccatcaaaaggagtgtatgtgcgaacggagtaaccacttcac 486  
 |||||

Db 189 GGAATAATACCCAAATGTGTTCAAGGCGTTTACGTTTCAAAACGGAGCTAACCCCTCTTTT 248  
 |||||  
 QY 487 gggccgggtgacagggtcacactcttcgcgagagacgggtatggttcacgcggcgaattc 546  
 |||||  
 Db 249 GAACCAACCGCCGGACACCATTTCTTCGACGGGACGGTATGGTTCCGCCGTTCAATTC 308  
 |||||  
 QY 547 gaacacgggttcagctagctacgcttgccgggtttactcagactaacccgggtttgttcaggaa 606  
 |||||  
 Db 309 AAAAATGGGTGGCTAGTTACGCTTGCCGTTTCACTGAAACAGAGAGGCTTTGTTCAAGAA 368  
 |||||  
 QY 607 cgtcaatgggtcgaccgggtttttcccaaaagccatcggtgagcttcacggccacacccgt 666  
 |||||  
 Db 369 AAAGCTTTGGGTGGCCCTGTTTCCCTAAAGCCATTGGTGAATTACATGCTCACTCTGGA 428  
 |||||  
 QY 667 attgcccgactcatgctattctacgcccagagctgcagccgggtatagtcgacccggcacac 726  
 |||||  
 Db 429 ATTGCAAGGCTTATGCTGTTTACGCTCGTGGGCTCTTCGGACTTGTGATCACAGTAA 488  
 |||||  
 QY 727 ggaacccgggtgtagctaacgcccgggtttggctctatttcaatggccgggtattggctatgtcg 786  
 |||||  
 Db 489 GGAACCTGGTGTTCGAAACGCCGCTTTAGTCTATTTCATTAACCGATTACTTGTATGTCT 548  
 |||||  
 QY 787 gaggatgatttacccttaccgaagttcagatcactcccaatggagatttaaaacccgttgg 846  
 |||||  
 Db 549 GAAGATGATTTCCTTACCCTTAAAGGTAAACCCACCGCGCATCTTAAACAGAAAGT 608  
 |||||  
 QY 847 cggttccgat 855  
 |||||  
 Db 609 CGATTCCGAT 617  
 |||||

Search completed: July 24, 2002, 06:53:47  
 Job time: 5972 sec

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: November 15, 2003, 09:15:54 ; Search time 6156 Seconds  
(without alignments)  
3980.648 Million cell updates/sec

Title: US-09-758-269-6  
Perfect score: 3150  
Sequence: 1 MASFTATAAASGRWLGNHT.....VPYGFHGTFIGADDLAKQVV 599

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2889711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2.1/USPTO.spool/US09758269/runat.14112003.135756.28745/app.query.fasta.1.775  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09758269 @CEN 1 1 3963 @runat.14112003.135756.28745 -NCPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.on.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*

29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.rtd.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3150	100.0	1800	6	AX148306 Sequence
2	3150	100.0	1800	6	BD017431 Transgeni
3	3150	100.0	3869	8	AB026549 Arabidops
C 4	3150	100.0	52232	8	AB028617 Arabidops
5	3143	99.8	2331	8	AY056255 Arabidops
6	2291	72.7	2164	8	STU276244 Solanum t
7	2282.5	72.5	2171	8	LENEOXANT
8	2282.5	72.5	19018	8	LES439079
9	2280.5	72.4	1818	6	AX148316 Sequence
10	2172	69.0	2398	8	BD017436 Transgeni
11	2170.5	68.9	95769	6	AF190462 Phaeolus
12	2170.5	68.9	95769	6	AF190462 Phaeolus
13	2170.5	68.9	95769	6	AC013430 Genomic s
14	2168.5	68.8	1839	6	AX148312 Sequence
15	2168.5	68.8	1839	6	BD017434 Transgeni
16	2168.5	68.8	2432	8	AB030293 Vigna ung
17	2155.5	68.4	2357	8	AB080192 Pisum sat
18	2075.5	65.9	2034	8	AB080193 Pisum sat
19	2066.5	65.6	2310	8	AF224671 Persea am
C 20	2044.5	64.9	72058	8	AC074176 Arabidops
21	1991	63.2	1752	6	AX148302 Sequence
22	1991	63.2	1752	6	BD017429 Transgeni
23	1991	63.2	110102	8	ATF28J12
24	1991	63.2	194143	8	ATCHRIV48
C 25	1988	63.1	175158	2	AC120531 Oryza sat
26	1988	63.1	181159	2	CNS071P7
27	1987	62.3	2153	8	AF224672 Persea am
28	1962	62.3	133983	8	CNS08CB9
29	1959.5	62.2	131689	8	CNS08CB9
30	1930	61.3	1815	6	AX148314 Sequence
31	1930	61.3	1815	6	BD017435 Transgeni
32	1930	61.3	2498	8	ZM095953 Zea mays v1
33	1834.5	58.2	149097	2	AP005632 Oryza sat
34	1663.5	52.8	1734	6	BD017430 Sequence
35	1663.5	52.8	1734	6	BD017433 Transgeni
36	1663.5	52.8	78921	8	AB028621 Arabidops
37	1214	38.5	82697	8	AT9A21
38	1014.5	32.2	120055	8	AC123596 Medicago
39	1003	31.8	80686	2	AC144759 Medicago
C 40	986.5	31.3	120055	8	AC123596 Medicago
41	986.5	31.3	2016	8	AB080194 Pisum sat
C 42	971.5	30.8	80686	2	AC144759 Medicago
43	569	30.8	1788	6	AX148304 Sequence
44	969	30.8	1788	6	AX506749 Sequence
45	969	30.8	1788	6	BD017430 Transgeni

ALIGNMENTS

RESULT 1

```

AX148306
LOCUS AX148306 1800 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 5 from Patent EP1116794.
ACCESSION AX148306
VERSION AX148306.1 GI:14347193
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
1
REFERENCE Iuchi, S., Kobayashi, M. and Shinozaki, K.
AUTHORS Transgenic plants carrying neoxanthin cleavage enzyme gene
TITLE Patent: EP 1116794-A 5 18-JUL-2001;
JOURNAL Riken (JP)
FEATURES
source
1..1800
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
1..1800
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC41194.1"
/db_xref="GI:14347194"
/translation="MASFTATAVSGRWLGNHTQPPLSQSSQSDLSYCSLPLMASRV
TRKLVNSALHTPPALHPKQSSNSPAIVVKPKAKESNTKOMNLFORAAAALDAAGS
FLVSEKHLPLKPTADPSVQIAGNAPVNEQPVRRNLFPVVGKLPDSIKGVYVRNGANP
LHEPVTGHHFPGDGMVHAVKFEHGSASYACRFQTNRFVQEROLGRPVPFKAI GEHL
GHTGIARMLFPYARAAGIIVDPAHGTGVANAGLVFNGRLLAMSEDDLPYQVOITPNG
DLKTVGRFDFDQLESFMIAHPKVDPSGELPALSYDVUSKPYLYKFRFPDPTKSPD
VEIQDQTMHDFATIEFNVFVDPQVVKLPEMIRGSGPVVDKMKVAFGLIDKY
AEDSSNTKIDAPDFCFHLMNAWEEPETDEVVVGISMTTPPDSIFNEDENLKSVL
EIRLNTKGTSTRRPIISNEDQVNLKGMVNRMLGRTKFAYIALAEWPVKVSGFA
KVDLTTEVKKHLYGDNRYGGEPLFLPGEGERDEGVILCFVHDEKTKWSELQIVNAV
SLEVEATKLPSPRYPGVPHGTFIFGADDLAKQVV"
BASE COUNT 458 a 464 c 439 g 439 t
ORIGIN
Alignment Scores:
Pred. No.: 6,73e-220 Length: 1800
Score: 3150.00 Matches: 599
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-758-269-6 (1-599) x AX148306 (1-1800)
Qy 1 MetAlaSerPheThrAlaThrAlaAlaValSerGlyArgTrpLeuGlyGlyAsnHisThr 20
Db 1 ATGGCTTCTTTTTCACGGCAACGGCTCGCGTTTTCGGGAGATGGCTTGGTGGCAATCATACT 60
Qy 21 GlnProLeuSerSerGlnSerSerAspLeuSerTyrCysSerSerLeuProMet 40
Db 61 CAGCGCCCATATCGTCTCTCTCAAGCTCCGACCTTGAGTTAGTCTCTTACCTATG 120
Qy 41 AlaSerArgValThrArgLysLeuAsnValSerSerAlaLeuHisThrProAlaLeu 60
Db 121 GCCAGTCGTGTACACAGTAAGCTCAATGTTTTCATCTGCGCTTTCACACTCTCTCCAGCTCTT 180
Qy 61 HisPheProLysGlnSerSerAsnSerProAlaLeuValValLysProLysAlaLysGlu 80
Db 181 CATTTCCCTAAGCAATCATCAAACTCTCCGCCCATTTGTTGTTAAGCCCAAGCAAGAA 240
Qy 81 SerAsnThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaLeuAspAlaAla 100
Db 241 TCCAACTAAACAGATGAATTTGTTCCAGAGACGGCGCGGCGGCTTGGACGGCGCG 300
Qy 101 GluGlyPheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSer 120
Db 301 GAGGTTTCTTGTGTGTCAGCCACGAGAGCTACACCGGCTTCTCTAAACCGGCTGATCTCTAGT 360

```

```
Qy 481 ValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeuAlaLeuAlaGluPro 500
Db 1441 GTCACAGAACATGCTCGCGCTAAACCAAAATTCGCTTACTTGCGCTTTAGCCGAGCG 1500
Qy 501 TrpProLysValSerGlyPheAlaLysValAspLeuThrGlyGluValLysLysHis 520
Db 1501 TGGCCTTAAGTCTCAGGATTCGCTAAGTGTGATCTCACTACTGGAGAAGTTAAGAAACAT 1560
Qy 521 LeuTyrGlyAspAsnArgTyrGlyGluProLeuPheLeuProGlyGluGlyGlu 540
Db 1561 CTTTACGGCGATAACCGTTACGGAGAGAGCCCTCTGTTCTCCCGGAGAGGAGAGAG 1620
Qy 541 GluAspGluGlyTyrLleLeuCysPheValHisAspGluLysThrTrpLysSerGluLeu 560
Db 1621 GAAGACGAGAGGATACATCCTCTGTTTCGTTTCACGACGAGACATGGAAATCGGAGTTA 1680
Qy 561 GlnIleValAsnAlaValSerLeuGluValGluAlaThrValLysLeuProSerArgVal 580
Db 1681 CAGATAGTTAAGCCGTTAGCTTAGAGGTTGAAGCAACGGTTAACTTCCGTCAGAGGTT 1740
Qy 581 ProTyrGlyPheHisGlyThrPheIleGlyAlaAspAspLeuAlaLysGlnValVal 599
Db 1741 CCGTACGGATTTACGGTACATTCATCGAGCGGATGATTTGGCGAAGCAGGTCGTG 1797

RESULT 2
BD017431 1800 bp DNA linear PAT 27-AUG-2002
LOCUS Transgenic plant using neozanthine cleaving enzyme gene.
DEFINITION
ACCESSION BD017431
VERSION BD017431.1 GI:22558607
KEYWORDS JP 2001258579-A/3.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 1800)
AUTHORS Iuchi, K., Kobayashi, M. and Shinozaki, K.
TITLE Transgenic plant using neozanthine cleaving enzyme gene
JOURNAL Patent: JP 2001258579-A 3 25-SEP-2001;
THE INSTITUTE OF PHYSICAL AND CHEMICAL RESEARCH
OS Arabidopsis thaliana (thale cress)
PN JP 2001258579-A/3
PD 25-SEP-2001
PF 11-JAN-2001 JP 2001003476
PI KIYOSHI IUCHI, MASATOMO KOBAYASHI, KAZUO SHINOZAKI PC
C12N15/09, A01H5/00, C12N5/02, C12N5/10, C12N15/00, C12N5/00 CC
Transgenic plant using neozanthine cleaving enzyme gene PH Key
Location/Qualifiers
FT CDS (1)..(1800).
Location/Qualifiers
source 1..1800
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"

BASE COUNT 458 a 464 c 439 g 439 t
ORIGIN

Alignment Scores:
Pred. No.: 6,738-220 Length: 1800
Score: 3150.00 Matches: 599
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-758-269-6 (1-599) x BD017431 (1-1800)

Qy 1 MetaLaserPheThrAlaThrAlaValSerGlyArgTrpLeuGlyGlyAsnHisThr 20
Db 1 ATGGCTTCTTTCACGGCAACGCTCGGTTCTCGGAGATGCGTTGGTGGCATCATACT 60
Qy 21 GlnProProLeuSerSerGlnSerSerAspLeuSerTyrCysSerSerLeuProMet 40
```

```
Db 61 CAGCGCGCATTTATCGTCTTCTCAAAGCTCCCAAGCTTGAAGTTATTTAGTCTCTTACCTATG 120
Qy 41 AlaSerArgValThrArgLysLeuAsnValSerSerAlaLeuHisThrProProAlaLeu 60
Db 121 GCCAGTCTGTGTACACGTAAGCTCAATGTTTCATCTCGGCTTTCACACTCTCTCCAGCTCTT 180
Qy 61 HisPheProLysGlnSerSerAsnSerProAlaIleValValLysProLysAlaLysGlu 80
Db 181 CATTTCCCTAAGCAATCATCAAACTCTCCCGCATTTGTTTAAAGCCCAAGCAAGAA 240
Qy 81 SerAsnThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaLeuAspAlaAla 100
Db 241 TCCAACACTAAACAGATGAATTTGTTCCAGAGAGCGCGCGGAGCGTTGGACCGCGC 300
Qy 101 GluGlyPheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSer 120
Db 301 GAGGGTTTCTTGTGTCACGACAGAGACTACCCGCTTCTCTAAACCGGCTGATCCTAGT 360
Qy 121 ValGlnIleAlaGlyAsnPheAlaProValAsnGluGlnProValArgAsnLeuPro 140
Db 361 GTTCAGATCGCGGAAATTTTGCTCCGTTGAATGAACAGCCGCTCCGCGCTAATCTTCG 420
Qy 141 ValValGlyLysLeuProAspSerLysLysGlyValTyrValArgAsnGlyAlaAsnPro 160
Db 421 GTGTCGGAACACTTCCGATTTCCATCAAGAGAGTGTATGTGCGCAACGGAGCTAACCCA 480
Qy 161 LeuHisGluProValThrGlyHisPhePheAspGlyAspGlyMetValHisAlaVal 180
Db 481 CTTTACGAGCCGGTGACAGGTCACCACTTCTTCAGCGAGAGCGGTATGGTTTCCGCCGTC 540
Qy 181 LysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheVal 200
Db 541 AAATTCGAACACGGTTTACGTAGCTTACGCTTTCGCGGTTTACTCAGACTAACCGGTTGTT 600
Qy 201 GlnGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGlyHis 220
Db 601 CAGGAACGTCATTTGGGTTCGACCGGTTTCCCAAGGCGCATTCGGTGAAGCTTCCAGCGCAC 660
Qy 221 ThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaAlaGlyIleValAspPro 240
Db 661 ACCGGTATTCGCCACTCATCTATTTCACGCCAGAGCTGCAGCCGCTATAGTTCAGCCCG 720
Qy 241 AlaHisGlyThrGlyValAlaAsnAlaGlyLeuValTyrPheAsnGlyArgLeuLeuAla 260
Db 721 GCACACGAACCGGTGTAGCTAACCGCGTTTGGTCTATTTCAATGGCCGGTTATTGGCT 780
Qy 261 MetSerGluAspAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysThr 280
Db 781 ATGTCGGAGGATGATTTTACCTTTACCAAGTTCCAGATCACTCCCAATGGAGATTTAAAAACC 840
Qy 281 ValGlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysVal 300
Db 841 GTTGGTCGTTTCGATTTTGTAGTGACAAATTAGATCCACAAATGATTGCCACCCGGAAGTC 900
Qy 301 AspProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValSerLysProTyrLeu 320
Db 901 GACCCGGAATCCGGTGAACCTCTTCGTTTAAAGCTACGACGCTCGTTTCAAGAGCTTTACCTA 960
Qy 321 LysTyrPheArgPheSerProAspGlyThrLysSerProAspValGluIleGlnLeuAsp 340
Db 961 AAATACTTCCGATTTCTACCGGACGGAACATAAATCACCGGACGCTCGAGATTTCAGTTGAT 1020
Qy 341 GlnProThrMetMetHisAspPheAlaIleThrGluAsnPheValValValProAspGln 360
Db 1021 CAGCCACGATGATGCACGATTTCCGATTTACAGAAACTTCGTCGTCTGCTACCTGACCCAG 1080
Qy 361 GlnValValPheLysLeuProGluMetIleArgGlyGlySerProValValTyrAspLys 380
Db 1081 CAAGTCGTTTCAAGCTCCCGGAGATGATCCGCGTGGGTCTCCGCTGGTTTACGACAAG 1140
Qy 381 AsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSerSerAsnIleLys 400
```

```
Db 1141 AACAGGTCGCAAGATTCGGGATTTTAGACAAATACGCCGAAGATTTCATCGAACATTAAAG 1200
Qy 401 TrpIleAspAlaProAspCysPheHisLeuTrpAsnAlaTrpGluProGlu 420
Db 1201 TGGATTGATGTCAGAGATTGCTTCTGCTTCATCTCTGGAACGCTTGGGAAGAGCCAGAA 1260
Qy 421 ThrAspGluValValIleGlySerCysMetThrProProAspSerIlePheAsnGlu 440
Db 1261 ACAGATGAAGTCGTCGTGATAGGTCCTGTATGACTCCACCAAGACTCAATTTTCAACGAG 1320
Qy 441 SerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsnLeuLysThrGlyGlu 460
Db 1321 TCTCAGCAGAACTCAAGAGTGTCTCTGTCGAAATCCGCTGAATCTCAAAACCGGTGAA 1380
Qy 461 SerThrArgArgProIleIleSerAsnGluAspGlnGlnValAsnLeuGluAlaGlyMet 480
Db 1381 TCAACTCGCCGTCGATCATCTCCACGAGATCAACAAGTCAACCTCGAAGCAGGAGATG 1440
Qy 481 ValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeuAlaLeuAlaGluPro 500
Db 1441 GTCAACAGAAACATGCTCGGCCGTAAACCAAAATTCGCTTACTTGGCTTTAGCCGAGCCG 1500
Qy 501 TrpProLysValSerGlyPheAlaLysValAspLeuThrThrGlyGluValLysLysHis 520
Db 1501 TGGCCTAAAGTCTCAGGATTCGCTAAGTTCATCTCACTACTCGAGAAGTTAAGAAACAT 1560
Qy 521 LeuTyrGlyAspAsnArgTyrGlyGluProLeuPheLeuProGlyGluGlyGlu 540
Db 1561 CTTTACGCGGATACCGGTACGGAGGAGAGCTCTGTCTTCTCCCGGAGAGGAGAGAG 1620
Qy 541 GluAspGluGlyTyrIleLeuCysPheValHisAspGluLysThrTrpLysSerGluLeu 560
Db 1621 GAAGCAGGAAGATACATCTCTGTTCGTTCCACGACGAGAGAGCATGAAATCGGAGTTA 1680
Qy 561 GlnIleValAsnAlaValSerLeuGluValGluAlaThrValLysLeuProSerArgVal 580
Db 1681 CAGATAGTTAACGCGGTAGCTTAGAGGTTGAAGCAACGGTTAACTTCGTCAGGGTT 1740
Qy 581 ProTyrGlyPheHisGlyThrPheIleGlyAlaAspLeuAlaLysGlnValVal 599
Db 1741 CCGTACGGAATTCACGGTACATTCATCGGACCGCATGATTTGGCGAAGCAGGTCGTG 1797

RESULT 3
AB026549 3869 bp DNA linear PLN 24-OCT-2001
DEFINITION Arabidopsis thaliana gene for neoxanthin cleavage enzyme, complete cds.
ACCESSION AB026549
VERSION AB026549.1 GI:16416373
KEYWORDS neoxanthin cleavage enzyme.
SOURCE Arabidopsis thaliana (chale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
REFERENCE Iuchi,S., Kobayashi,M. and Shinozaki,K.
AUTHORS Characterization of neoxanthin cleavage enzyme from Arabidopsis
TITLE thaliana
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3869)
AUTHORS Iuchi,S. and Shinozaki,K.
TITLE Direct Submission
JOURNAL Submitted (22-APR-1999) Satoshi Iuchi, RIKEN, Plant Mol. Bio.;
3-1-1 Koyiyadai, Tsukuba 305-0074, Japan
(E-mail:iuchi@rtc.riken.go.jp, Tel:81-298-36-4359)
FEATURES
Location/Qualifiers
1..3869
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Col."
/db_xref="taxon:3702"
1831..3630
CDS
```

```
/codon_start=1
/product="neoxanthin cleavage enzyme"
/protein_id="BAB70609.1"
/db_xref="GI:16416374"
/translation="MASFTATAAASVGRWLGNNHTQPLSSQSSDLSYCSLPMASRV
TRLNVSALHTPPALFPKQSSNSPAIVVKPAKESNTKNNLFORAAAAALDAAG
FLVSHKLPKPTADPSVQIAGNFAPVNEQPVRRNLPVVVKLPDSTKGVYVRNGANP
LHEPVTGHHFFDGDGMVHAKVPEHGSASVACRFTQTNRFVQERQLGRPVPPKAIIEUH
GHTGIARLMLFYARAAAGLIVPAHGTGVANAGLVYFNGRLANSEDDLPYQVQITPNG
DLTKGRFDFDQLESTMIAPKVPDSEGFALSDVSVKLPKPEMIRGSSPVYDKNNKVARFGLDKY
VETQLQPTMMHDAITENFVVVDQVVFVKLPEMIRGSSPVYDKNNKVARFGLDKY
AEDSNIKWIDAPDCFCFLMNAWEPETDEVVIVGSCMTTPDPSINESDENKLSVLS
EIRLNLTGEBSTRPPIISNEDQOONLEAGMVRNMLGRKTFAYLALAEWPVKVSGFA
KVDLTGGEVKKHLYGDNRYGGEPLFLPGEGGEDEGYILLFVHDEKTKMSELQIVNAV
SLSEVATVKLPSRVPVGFHGTGFIADDLAKQVV"
BASE COUNT 1155 a 845 c 773 g 1096 t
ORIGIN
Alignment Scores:
Pred. No.: 1,99e-219 Length: 3869
Score: 3150.00 Matches: 599
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0
US-09-758-269-6 (1-599) x AB026549 (1-3869)
Qy 1 MetAlaSerPheThrAlaThrAlaAlaValSerGlyArgTrpLeuGlyGlyAsnHisThr 20
Db 1831 ATGGCTTCTTTCACGGCAACGGCTGGCTTCTGGGAGATGGCTTGGTGGCAATCATACT 1890
Qy 21 GlnProLeuSerSerSerGlnSerSerAspLeuSerTyrCysSerSerLeuProMet 40
Db 1891 CAGCGCCAVTATGCTTCTCAAAGCTCCGACTTGAGTTATTGTAGCTCTCTTACCTATG 1950
Qy 41 AlaSerArgValThrArgLysLeuAsnValSerAlaLeuHisThrProAlaLeu 60
Db 1951 GCCAGTCGTGTACACGTAAGCTCAATGTTTCACTCGGCTTCACACTCTCTCCAGCTCTT 2010
Qy 61 HisPheProLysGlnSerSerAsnSerProAlaIleValValLysProLysAlaLysGlu 80
Db 2011 CATTTCCCTAAGCAATCATCAAACTCTCCGCCATTGTTGTTAAGCCCAAGCAAGAA 2070
Qy 81 SerAsnThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaLeuAspAlaAla 100
Db 2071 TCCAACACTAAACAGATGAATTTGTTCCAGAGAGCGCGCGGCGGCGGCGGCG 2130
Qy 101 GluGlyPheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSer 120
Db 2131 GAGGGTTTCTTGTACCCACGAGAGACTACACCCGCTTCTTAAACAGGCTGATCTTAGT 2190
Qy 121 ValGlnIleAlaGlyAsnPheAlaProValAsnGluGlnProValArgArgAsnLeuPro 140
Db 2191 GTTCAGATCGCCGAAATTTTGTCCGCGTGAATGAACAGCCCGTCCGCGGCTAATCTCCG 2250
Qy 141 ValValGlyLysLeuProAspSerIleGlyGlyValTyrValArgAsnGlyAlaAsnPro 160
Db 2251 GTGGTCGAAAACTTCCCGATTCCATCAAGAGAGTGTATGTGCGCAACGCGAGCTTAAAC 2310
Qy 161 LeuHisGluProValThrGlyHisHisPhePheAspGlyAspGlyMetValHisAlaVal 180
Db 2311 CTTCCAGAGCGGCGGAGAGGTCACCACTTCTTCACGAGAGCGGATGTTGTTTACCGCGTC 2370
Qy 181 LysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheVal 200
Db 2371 AAATTCGAACACACGGTTTTCAGCTAGCTACGCTTTCGCGGTTTACTCAGACTTAACCGGTTGTT 2430
Qy 201 GlnGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGlyHis 220
Db 2431 CAGGAACGTCAATTTGGGTGACCGGTTTTCCTCCAAAGCCATCGGTGAGCTTCACGGCCAC 2490
Qy 221 ThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaGlyIleValAspPro 240
```

```
Db 2491 ACCGGTATTCGGCGACTCATGTATTCTACGCGAGAGCTGCAGCCGGTATAGTCGACCCG 2550
Qy 241 AlaHisGlyThrGlyValAlaAsnAlaGlyLeuValThrPheAsnGlyArgLeuLeuAla 260
Db 2551 GCACACGGAACCGGTAGTACGCGCGTGTGGTCTATTCTATTTCAATGGCCGGTATTGGCT 2610
Qy 261 MetSerGluAspAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysThr 280
Db 2611 ATGTCCGAGGATGATTACCTTACCAGAGTTCAGATCACTCCCAATGGAGATTAAACCC 2670
Qy 281 ValGlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysVal 300
Db 2671 GTTGTGCGTTTGATTTTATGATGACAAATTAGAAATCACAATGATTCGCCACCCCGAAGTC 2730
Qy 301 AspProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValSerLysProTyrLeu 320
Db 2731 GACCCGGAATCCGGTGAACCTCTTCGTTTAAGTACGAGCTGCTTCAAGGCTTACCTA 2790
Qy 321 LysTyrPheArgPheSerProAspGlyThrLysSerProAspValGluIleGlnLeuAsp 340
Db 2791 AAATACTTCCGATTCCTCCGCGACGGAATAAATCACCGAGCTCGAGATTCAGCTTGAT 2850
Qy 341 GlnProThrMetMetHisAspPheAlaIleThrGluAsnPheValValProAspGln 360
Db 2851 CAGCCAACGATGATCACGATTTCCGATTTACAGAGAACTTCGTCTGCTGATCTGACCTGAC 2910
Qy 361 GlnValValPheLysLeuProGluMetIleArgGlyGlySerProValValTyrAspLys 380
Db 2911 CAAGTCGTTTCAAGCTCCGAGATGATCCGCGGTGGTCTCCGGTGGTTTACGACAA 2970
Qy 381 AsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSerSerAsnIleLys 400
Db 2971 AACAAAGGTCCGAAGATTCGGGATTTAGACAAATACCGCGAAGATTCATCGAACATTA 3030
Qy 401 TrpIleAspAlaProAspCysPhePheHisLeuTrpAsnAlaTrpGluLeuProGlu 420
Db 3031 TGGATTGATGCTCCAGATTGCTTCTGCTCCATCTCTGGAACGCTTGGGAAGCGCCAGAA 3090
Qy 421 ThrAspGluValValIleGlySerCysMetThrProProAspSerIlePheAsnGlu 440
Db 3091 ACAGATGAAGTCGTCGTATAGGTCCTGTATGACTCCACACAGATCAATTTTCAACGAG 3150
Qy 441 SerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsnLeuLysThrGlyGlu 460
Db 3151 TCTGACGAGAATCTCAAGAGTCTCTGTCTGAATCCGCTGAAATCTCAAAACCGGTGAA 3210
Qy 461 SerThrArgArgProIleIleSerAsnGluAspGlnGlnValAsnLeuGluAlaGlyMet 480
Db 3211 TCAACTCCCGCTCCGATCATCTCCAACGAAGATCAACAAGTCAACCTCGAAGCAGGGATG 3270
Qy 481 ValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeuAlaLeuAlaGluPro 500
Db 3271 GTCAACAGAAACATGCTCGCCGCTAAACCAATTCGTTACTTGGCTTTAGCCGAGCG 3330
Qy 501 TrpProLysValSerGlyPheAlaLysValAspLeuThrThrGlyGluValLysLysHis 520
Db 3331 TGGCCCTAAAGTCTCAGGATTCCTAAAGTTGATCTCACTACTGGAGAAGTTAAGAAACAT 3390
Qy 521 LeuTyrGlyAspAsnArgTyrGlyGluProLeuPheLeuProGlyGluGlyGlyGlu 540
Db 3391 CTTTACGGCGATTAACCGTTTACGGAGGAGAGCCTCTGTTTCTCCCGGAGAAGGAGAG 3450
Qy 541 GluAspGluGlyTyrIleLeuCysPheValHisAspGluLysThrTrpLysSerGluLeu 560
Db 3451 GAAGACGAGGATACATCTCTGTTTCTGTTTACGACGAGAGACATGAAATCGGAGTTA 3510
Qy 561 GlnIleValAsnAlaValSerLeuGluValGluAlaThrValLysLeuProSerArgVal 580
Db 3511 CAGATAGTTAACCGCTTAGCTTAGAGGTTGAAGCAACGGTTAAACTTCCGTCAGGGTT 3570
Qy 581 ProTyrGlyPheHisGlyThrPheIleGlyAlaAspAspLeuAlaLysGlnVal 599
```

```
Db 3571 CCGTACGATTCACGGTACATTCTATCGGACCGGATGATTTGGCGAAGCAGGTCTGTG 3627
RESULT 4
AB028617/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (sites)
Sato,S., Nakamura,Y., Kaneko,T., Kato,T., Asamizu,E. and Tabata,S.
Structural analysis of Arabidopsis thaliana chromosome 3. I.
Sequence features of the regions of 4,504,864 bp covered by sixty
P1 and TAC clones
DNA Res. 7 (2), 131-135 (2000)
20277480
10819329
2 (bases 1 to 52232)
Sato,S., Nakamura,Y., Kaneko,T., Kato,T., Asamizu,E. and Tabata,S.
Submitted (09-JUN-1999) Yasukazu Nakamura, Kazusa DNA Research
Institute, Department of Plant Gene Research; 1532-3, Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail: ynakamu@kazusa.or.jp,
Tel: 81-438-52-3935, Fax: 81-438-52-3934)
Address for correspondence: kaos@kazusa.or.jp
For the latest information on annotation of this clone, please see
http://www.kazusa.or.jp/kaos/cgi-bin/agd/graph.cgi?c=MOA2
Genes with similarity to proteins in the databases are described in
'product' or 'note' qualifiers. Genes that have no significant
protein similarity are described as 'unknown protein'.
The software programs used to predict genes include: Grail
(Informatics Group, Oak Ridge National Laboratory,
http://compbio.ornl.gov/Grail-1.3/),
GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
SplicePredictor (Volker Brendel, Stanford University,
http://gremlini.zoology.ualb.ca/edu/cgi-bin/sp.cgi).
Genes encoding tRNAs are predicted by tRNAscan-SE
(Sean Eddy, Washington University School of Medicine, St. Louis,
http://genome.wustl.edu/eddy/tRNAscan-SE/).
This sequence may not be the entire insert of this clone. It may be
shorter because we remove overlaps between neighboring submissions.
The 5' clone is MLN21 and the 3' clone is MIE1.
FEATURES
Location/Qualifiers
1. 52232
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="3"
/clone="MOA2"
/clone_lib="Mitsui PI"
join(1596..2132,2225..2349,2455..2501,2648..2725,
2807..2931,3021..3087,3216..3278,3370..3439,3529..3629,
3718..3822,3906..4091)
/note="gene_id:MOA2.2"
/codon_start=1
/evidence=not experimental
/product="glycolate oxidase"
/protein_id="BAB0134.1"
/db_xref="GI:1194212"
/translat="MEITNVTEYDAIAKQKPKMYDYVYASGAEDQWTLQENRNARFAR
ILFPRILIDSVKIDMTTTLVGLKISMPIMVAPTAKMAHPDGEYATARAASAAGTI
MTLSSWATSSVEEAVATPGPIRFQLYVYKRNVRVQQLVRAERAGFAKIALTVDPTR
LGRSEDIKRPFTLPNLTAKNFEGLDGLKMDKDEANDSLASYAGQIDRTLSWKDVQV
LQITKLPILVKVGTEDARIALQAGAGGIIVSNHNGARLDYVPATISALEEVVKAT
QGRIPVFLDGGVRRGTDFVKALALGASGIFIGRPVVFSLAAEGEAGVRKVLQMLRDEF
```





QY 21 GlnProProLeuSerSerSerGlnSerSerAspLeuSerTyrCysSerSerLeuProMet 40  
DB 13609 CAGCGCCCAATATCGCTCTTCAAAAGCTCCGACTTGGATTATTTAGCTCTTACCTATG 13550  
QY 41 AlaSerArgValThrArgLysLeuAsnValSerSerAlaLeuHisThrProProAlaLeu 60  
DB 13549 GCCAGTCGTGTACACAGTAAGCTCAATGTTTCATCTGCGCTTCACTCTCTCAGCTCTT 13490  
QY 61 HisPheProLysGlnSerSerSerSerProAlaIleValValProLysAlaLysGlu 80  
DB 13489 CATTTCCCTTAAGCAATCATCAAACTCTCCCGCAATTTGTTAAGCCCAAGCAAGAA 13430  
QY 81 SerAsnThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaAlaLeuAspAlaAla 100  
DB 13429 TCCAACTAAACACAGATGAATTTGTTCCAGAGAGCGCGCGCGCTGGACCGCGG 13370  
QY 101 GluGlyPheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSer 120  
DB 13369 GAGGGTTTCTTGTGAGCCACGAGAGCTACACCGCTTCTTAAACGGCTGATCCTAGT 13310  
QY 121 ValGlnIleAlaGlyAsnPheAlaProValAsnGlnProValArgAsnLeuPro 140  
DB 13309 GTTCAGATCGCGGAAATTTGCTCCGGTGAATGAACAGCCCGTCCGGCTAATCTCCG 13250  
QY 141 ValValGlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyAlaAsnPro 160  
DB 13249 GTGGTCGGAAATCTTCCGATTTCCATCAAGAGGTATGTGCGCAACGGAGCTAACCCA 13190  
QY 161 LeuHisGluProValThrGlyHisPhePheAspGlyAspGlyMetValHisAlaVal 180  
DB 13189 CTTACAGAGCGGTGACAGGTACACCTTCTTCGACGAGAGCGTATGTTTACCGCGTC 13130  
QY 181 LysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheVal 200  
DB 13129 AAATTCGAACACGGTTACGTAGCTACGCTCTCCCGGTTTACTCAGACTAACCGGTTTGT 13070  
QY 201 GlnGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGlyHis 220  
DB 13069 CAGGAACGTCAATTTGGGTGACCGGTTTTTCCCAAGGCCATCGGTGAGCTTCAACGCCAC 13010  
QY 221 ThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaAlaGlyIleValAspPro 240  
DB 13009 ACCGGTATTTGCCGACTCATGCTATTCTACGCCAGAGCTGACGCCGTTATGTCAGCCCG 12950  
QY 241 AlaHisGlyThrGlyValAlaAsnAlaGlyLeuValTyrPheAsnGlyArgLeuLeuAla 260  
DB 12949 GCACACGGAACCGGTGAGTAAACCGGTTTGGTCTATTTCATATGGCCGGTATTGGCT 12890  
QY 261 MetSerGluAspAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysThr 280  
DB 12889 ATGTCGGAGGATGATTTTACCTTACCAAGTTTCAGATCACTCCCAATGGAGATTTAAAAACC 12830  
QY 281 ValGlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysVal 300  
DB 12829 GTTGTTCGGTTTCGATTTTGTGAGCAATAGTAATCCAAATGATGGCCACCCGGAAGTC 12770  
QY 301 AspProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValSerLysPheProTyrLeu 320  
DB 12769 GACCCGGAATCCGGTGAATCTTCTGCTTAACTAGCTACGAGCTGTTTCAAGCCTTACCTA 12710  
QY 321 LysTyrPheArgPheSerProAspGlyThrLysSerProAspValGluIleGlnLeuAsp 340  
DB 12709 AAATACTTCCGATTTCTACCGGACGGAATAATCACCAGGAGCTCGAGATTCAGCTTCAT 12650  
QY 341 GlnProThrMetMetHisAspPheAlaIleThrGluAsnPheValValProAspGln 360  
DB 12649 CAGCCAAACGATGATGACGATTTCCGATTAACAGAACTTCGTCTGCTGATTCACCGAG 12590  
QY 361 GlnValValPheLysLeuProGluMetIleArgGlyGlySerProValValTyrAspLys 380  
DB 12589 CAAGTCGTTTCAAGCTCCCGAGATGATCCCGGTGGGTCTCCGGTGGTTTACGACAG 12530  
QY 381 AsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSerSerAsnIleLys 400

DB 12529 AACAAAGGTTCGAAGATTTCGGATTTTAGACAAATACGCGGAAGATTTCATCGAACATTAAG 12470  
QY 401 TrpIleAspAlaProAspCysPheHisLeuTyrAsnAlaTyrGluGluProGlu 420  
DB 12469 TGGATTGATGTCACAGATTGCTTCTGCTTCAATCTCTGGAAGCGCTTGGGAAGAGCCAGAA 12410  
QY 421 ThrAspGluValValValIleGlySerCysMetThrProProAspSerIlePheAsnGlu 440  
DB 12409 ACAGATGAAGTCGTCGTATAGGGTCTGTATGACTCCACAGACTCAATTTTCAACGAG 12350  
QY 441 SerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsnLeuLysThrGlyGlu 460  
DB 12349 TCTGACGAGAATCTCAAGAGTGTCTGTCTGAAATCCGCTGAATCTCAAAACCGGTGAA 12290  
QY 461 SerThrArgArgProIleIleSerAsnGluAspGlnGlnValAsnLeuGluAlaGlyMet 480  
DB 12289 TCACTCGCCGTCGATCATCTCCAAACGAAATCAACAACTCACTCGAACGAGGATG 12230  
QY 481 ValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeuAlaLeuAlaGluPro 500  
DB 12229 GTCAACAGAAACATGCTCGCGCGTAAACCAAAATTCGCTTACTTGGCTTTAGCCGAGCGG 12170  
QY 501 TrpProLysValSerGlyPheAlaLysValAspLeuThrThrGlyGluValLysLysHis 520  
DB 12169 TGGCTTAAGTCTCAGGATTCGTAAGTTGATCTCTACTTGGAGAAAGTTAAGAAACAT 12110  
QY 521 LeuTyrGlyAspAsnArgTyrGlyGluProLeuPheLeuProGlyGluGlyGlu 540  
DB 12109 CTTTACGCGATACCGTTACGAGGAGGAGCTCTGTTTCTCCCGAAGAGGAGAGAG 12050  
QY 541 GluAspGluGlyTyrIleLeuCysPheValHisAspGluLysThrTrpLysSerGluLeu 560  
DB 12049 GAACAGCAAGGATACATCTCTGTTTCGTTCCACGAGAGAGACATGGAATTCGAGTTA 11990  
QY 561 GlnIleValAsnAlaValSerLeuGluValGluAlaThrValLysLeuProSerArgVal 580  
DB 11989 CAGATAGTTAACGCGCTTAGCTTAGAGTTGAAGCAACGTTAAACTTCCGTCAGGGTT 11930  
QY 581 ProTyrGlyPheHisGlyThrPheIleGlyAlaAspLeuAlaLysGlnValVal 599  
DB 11929 CCGTACGATTTACGGTACATTCATCGAGCCCGATGATTTGGCGAAGCAGGTCGTG 11873

## RESULT 5

LOCUS AY056255 2331 bp mRNA linear PLN 18-SEP-2002  
DEFINITION Arabidopsis thaliana putative 9-cis-epoxycarotenoid dioxygenase  
(At3g14440) mRNA, complete cds.  
ACCESSION AY056255  
VERSION AY056255.1 GI:15810432  
KEYWORDS FLI CDNA.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi  
1 (bases 1 to 2331)  
Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K.,  
Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L.,  
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,  
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,  
Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,  
Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K.,  
Davis, R.W., Ecker, J.R. and Theologis, A.  
Arabidopsis Full Length cDNA Clones  
Unpublished  
2 (bases 1 to 2331)  
Yamada, K., Banh, J., Banno, F., Dale, J.M., Goldsmith, A.D., Lee, J.M.,  
Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Yamamura, Y.,  
Yu, G., Yu, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R.,  
Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G.,  
Kawai, J., Kim, C., Koesema, E., Lam, B., Lin, J., Meyers, M.C.,  
Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,

Satou,M., Seki,M., Shinn,P., Southwick,A., Tracy,S.E.,  
 Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.  
 Direct Submission  
 Submitted (12-SEP-2001) Plant Gene Expression Center, 800 Buchanan  
 Street, Albany, CA 94710, USA  
 RIKEN Genomic Sciences Center (GSC) members carried out the  
 collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN  
 Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J.,  
 Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,  
 Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGSC (SSP) Consortium members carried out the  
 sequencing and annotation of the RAFL cDNAs: Yamada,K., Banno,J.,  
 Banno,F., Dale,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S.,  
 Quach,H.L., Tang,C.C., Toriumi,M., Yamamura,Y., Yu,G., Yu,S.,  
 Bowser,L., Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C.,  
 Koesema,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Nguyen,M.,  
 Palm,C.J., Shinn,P., Southwick,A., Tracy,S.E., Davis,R.W.,  
 Ecker,J.R. and Theologis,A.

Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to  
 this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC)  
 contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis  
 genome submitted to GenBank.

#### FEATURES

source  
 Location/Qualifiers  
 1. .2331  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /chromosome="3"  
 /clone="RAFL08-11-H16 (R11094)"  
 /note="This clone is in a modified pBluescript vector  
 (FLC-1) as a BamHI/XhoI insert."  
 ecotype: Columbia  
 1. .2331  
 /gene="At3g14440"  
 1. .122  
 /gene="At3g14440"  
 123. .1922  
 /gene="At3g14440"  
 /codon\_start=1  
 /evidence=experimental  
 /product="putative 9-cis-epoxycarotenoid dioxygenase"  
 /protein\_id="AA107104.1"  
 /translation="MASFTATAVSGRWLGNGHTQPPPLSSSQSSDLSCSLPMASRV  
 TRKNVSSALHTPPALHPKQSSNPAIVVKPKAKESNTKQMLFORAAALDAARG  
 FLVSHKHLPLPKTADPSVOLAGNFAPVNEQPVPRNLVVVKLPDSIKGVYVRGANGP  
 LHEPVTHHFPDGGGMVHAKFEGHSAYACRFQTRNFVQERQLRPVFPKALGELH  
 GHTGIARMLUFYABAAGIADVPKAGTGVANAGLVYFNGRLIAMSDDLQYQVQITPG  
 DLKTVGRFDGQLESTMIAPKVPDSEGLFALSDVSKPYLKYPFRFPDGTGKSPD  
 VEIQLDQPTWMHDAITENFVVDPQOVVFKLPEMILGSGSPVVDKKNKVARFGILDKY  
 AEDSNTKIDAPDFCFHLMWEEPEDEVVVIGSCMTPPDSIFNESDENLKSVL  
 EIRLNLTGSTRPIISNEDQVNLKAGMVRNMLGRKTKFAYLALAEWPVKVSGFA  
 KVDLTGEVRKRLHGDNRNYGGEPUFLFEGGEDEGILFCVHDEKTKWSELQILVNAV  
 SLEAVATVKLPSRVPYGFHGTFIGADLAKQVV"  
 1234  
 /misc\_difference  
 /gene="At3g14440"  
 /note="compared to genomic sequence resulting in an amino  
 acid sequence difference"  
 /replace="g"  
 1923. .2331  
 /gene="At3g14440"  
 2059  
 /misc\_difference  
 /gene="At3g14440"  
 /note="not present in genomic sequence"  
 2312  
 /misc\_difference  
 /gene="At3g14440"  
 /note="not present in genomic sequence"  
 2315  
 /misc\_difference  
 /gene="At3g14440"  
 1022

BASE COUNT 615 a 550 c 533 g 623 t  
 ORIGIN

#### Alignment Scores:

Pred. No.: 3.14e-219 Length: 2331  
 Score: 3143.00 Matches: 598  
 Percent Similarity: 99.83% Conservative: 0  
 Best Local Similarity: 99.83% Mismatches: 1  
 Query Match: 99.78% Indels: 0  
 DB: 8 Gaps: 0

US-09-758-269-6 (1-599) x AY056255 (1-2331)

Qy 1 MetAlaserPheThrAlaThrAlaAlaValSerGlyArgTrpLeuGlyGlyAsnHisThr 20  
 Db 123 ATGGCTCTTTTCACGGCAACGGCTCGGTTTCTGGGAGATGGCTTGGTGGCAATCATACT 182  
 Qy 21 GlnProProLeuSerSerSerGlnSerSerAspLeuSerTyrCysSerSerLeuProMet 40  
 Db 183 CAGCGCGCATATTCGTTCTCTCAAGCTCCGACTTGAGTTATTGTAGCTCTCTTACCTATG 242  
 Qy 41 AlaserArgValThrArgLysLeuAsnValSerSerAlaLeuHisThrProProAlaLeu 60  
 Db 243 GCCAGTCGTGTACACGTAAGCTCAATGTTTCACTCGCTTCACACTCTCTCCAGCTCTT 302  
 Qy 61 HisPheProLysGlnSerSerAsnSerProAlaIleValLysProLysAlaLysGlu 80  
 Db 303 CATTTCCCTAAGCAATCATCAAACTCTCCGCCATTGTTTAAAGCCCAAGGCCAAAGAA 362  
 Qy 81 SerAsnThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaAlaLeuAspAlaAla 100  
 Db 363 TCCNACACTTAACAGATGAATTTGTTCCAGAGAGCGCGCGGAGGTTGGAGCGCGCG 422  
 Qy 101 GluGlyPheLeuValSerHisGluLysLeuHiProLeuProLysThrAlaAspProSer 120  
 Db 423 GAGGGTTTCCTTGTCAACCCACGAGAGTACACCGCTTCTCTAAACGGCTGATCTCTAGT 482  
 Qy 121 ValGlnIleAlaGlyAsnPheAlaProValAsnGluGlnProValArgArgAsnLeuPro 140  
 Db 483 GTTCAGATCCCGGAAATTTTGCTCCGGTGAATGAACAGCCCGCTCCGGCGGTAACTTCCG 542  
 Qy 141 ValValGlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyValAsnPro 160  
 Db 543 GTGGTCGAAACACTTCCCGATTCCATCAAGAGAGTGTATGTGCGCAACGGAGCTAACCCA 602  
 Qy 161 LeuHisGluProValThrGlyHisHisPhePheAspGlyAspGlyMetValHisAlaVal 180  
 Db 603 CTTCCAGCGCGGTGACAGGTCACCACTTCTTCAGCGAGACGGTATGGTTTACGCGGTC 662  
 Qy 181 LysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheVal 200  
 Db 663 AAATTCCGAACACGGTTCAGCTAGCTACGCTTGCCTGCGGTTTACTCAGACTAACCGGTTTGT 722  
 Qy 201 GlnGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGlyHis 220  
 Db 723 CAGNAGCTCAATTTGGTGCACCGGTTTCCCAACGCCATCGGTGAGCTTCACGGCCAC 782  
 Qy 221 ThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaAlaGlyIleValAspPro 240  
 Db 783 ACCGGTATTGCCGACTCATGCTATTCTACGCCAGAGCTGCAGCGGTATAGTCGACCCG 842  
 Qy 241 AlaHisGlyThrGlyValAlaAsnAlaGlyLeuValTyrPheAsnGlyArgLeuLeuAla 260  
 Db 843 GCACACGAAACCGGTGTAGTAAACCGCGGTTTGGTCTTCTTCAATATGCGCGGTATTGGCT 902  
 Qy 261 MetSerGluAspAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysThr 280  
 Db 903 ATGTCGGAGGATGATTACCTTACCAAGTTCCAGATCACTCCCAATGGAGATTTTAAACACC 962  
 Qy 281 ValGlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysVal 300  
 Db 963 GTTGGTCGGTTCCGATTGTTGATGGACAATTAGAAATCCCAATGATTGTCGCCACCCGAAAGTC 1022

QY 301 AspProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValSerLysProTyrLeu 320  
|||  
Db 1023 GACCCGGAATCGGTTGAACCTCTCGCTTTAAAGTACGACGCTGTTCAAAGCCTTACCTA 1082  
|||  
QY 321 LysTyrPheArgPheSerProAspGlyThrLysSerProAspValGluLeuGlnLeuAsp 340  
|||  
Db 1083 AAATACTTCGGATTCTCACCAGCAGCACTAAATCACCAGCGCTGAGATTGAGCTTGAT 1142  
|||  
QY 341 GlnProThrMetHisAspPheAlaLeuThrGluAsnPheValValProAspGln 360  
|||  
Db 1143 CAGCCAAACGATGATGACGATTCGCGATTACAGAGAACTTCGTCGTCTACTCGACAG 1202  
|||  
QY 361 GlnValValPheLysLeuProGluMetLleArgGlyGlySerProValValTyrAspLys 380  
|||  
Db 1203 CAAGTCGTTTCAAGTCGCGAGATGATCCTCGGTGGTCTCCGGTGGTTTACGACAAG 1262  
|||  
QY 381 AsnLysValAlaArgPheGlyLleLeuAspLysTyrAlaGluAspSerSerAsnLys 400  
|||  
Db 1263 AACAGGTCGCAAGATTCGGGATTTTAGCAATACGCCGAAGATTCATCGAACATTAAG 1322  
|||  
QY 401 TrpLleAspAlaProAspCysPheCysPheHisLeuTrpAsnAlaTrpGluGluProGlu 420  
|||  
Db 1323 TGGATTGATGCTCCAGATTGCTTCTGCTTCATCTCTGGAACGCTTGGGAAGAGCAGNA 1382  
|||  
QY 421 ThrAspGluValValLleGlySerCysMetThrProProAspSerLlePheAsnGlu 440  
|||  
Db 1383 ACAGATGAAGTCGTCGATAGGTCCTGTATGACTCCACACAGACTCAATTTTCAACGAG 1442  
|||  
QY 441 SerAspGluAsnLeuLysSerValLeuSerGluLleArgLeuAsnLeuLysThrGlyGlu 460  
|||  
Db 1443 CTCGACGAGAACTCAAGAGTGTCTGTCTGAAATCCGCTCAATCTCAAAACCGGTGAA 1502  
|||  
QY 461 SerThrArgArgProLleSerAsnGluAspGlnGlnValAsnLeuGluAlaGlyMet 480  
|||  
Db 1503 TCAACTCGCGTCCGATCATCTCCACGAGAGTCAACAGTCAACTCGAAGCAGGATG 1562  
|||  
QY 481 ValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeuAlaLeuAlaGluPro 500  
|||  
Db 1563 GTCAACAGAAATGCTCGCGCTAAACCAAAATTCGCTTACTTGGCTTTAGCCGAGCGG 1622  
|||  
QY 501 TrpProLysValSerGlyPheAlaLysValAspLeuThrGlyGluValLysLysHis 520  
|||  
Db 1623 TGGCTTAAGTCTCAGGATTCCTAAAGTTGATCTCACTACTGGAAGAGTTAAGAAACAT 1682  
|||  
QY 521 LeuTyrGlyAspAsnArgTyrGlyGluProLeuPheLeuProGlyGluGlyGlu 540  
|||  
Db 1683 CTTTACGGGATACCGTTACGAGAGAGCGCTCTGTTTCTCCCGGAGAGGAGAGAG 1742  
|||  
QY 541 GluAspGluGlyTyrLleLeuCysPheValHisAspGluLysThrTrpLysSerGluLeu 560  
|||  
Db 1743 GAAGACGAAGGATACATCTCTGTTTCGTTTACGACGAGACACATGGAATCGAGTTA 1802  
|||  
QY 561 GlnLleValAsnAlaValSerLeuGluValGluAlaThrValLysLeuProSerArgVal 580  
|||  
Db 1803 CAGATAGTTAACCGCTTAGCTTAGAGTTGAAGCAACGGTTTAAACTTCCCGTCAAGGTT 1862  
|||  
QY 581 ProTyrGlyPheHisGlyThrPheLleGlyAlaAspAspLeuAlaLysGlnValVal 599  
|||  
Db 1863 CCGTACGGAATTCACGGTACATTCATCGGAGCGGATGATTTGGCAAGCAGGTCTGTG 1919  
|||

RESULT 6  
STU276244  
LOCUS  
DEFINITION  
Solanum tuberosum mRNA for putative 9-cis-epoxycarotenoid  
dioxygenase (nced1 gene).  
ACCESSION  
AJ276244  
VERSION  
9-cis-epoxycarotenoid dioxygenase; nced1 gene.  
KEYWORDS  
Solanum tuberosum (potato)  
SOURCE  
Solanum tuberosum  
ORGANISM  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
PUBLISHED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
1..2164  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4113"  
1..2164  
/gene="nced1"  
73..1887  
/gene="nced1"  
/function="cleavage of 9-cis-epoxycarotenoids"  
/codon\_start=1  
/product="putative 9-cis-epoxycarotenoid dioxygenase"  
/protein\_id="CAB76920.1"  
/db\_xref="GI:7209269"  
/translation="MATTTSHTNTWIKPLSPFSSKEFGFASNSISLLKQHNQRL  
NINSLOAPPILHPKOSNNYQTKPTSTISHPKQNNSSSSISKNVLQVAAAALD  
AVEGALTGHELEHLPKTPADPRVOISGNFAPVPENPCQSLPTVKIKPCVGVVVRN  
GANLPFETAGRHFDGDMVHAFQKNGSASACRFTETEFVQEKALGRPVFKAI  
GELHSGIARLMLFYARGLFGLDHSRGTVANGLVYFNNRLLMSEDDLLPHVKV  
TPTGDLTGREFDFDQKSTIAHPKLDPSGELFALSVDIOPKYLKRYFRSKNGE  
KSNDEIPEVEDPTMMHDFAI TEKFVII PQQVFMSEMRGGSFVVDYKNSRFGI  
LDKAKDGSDLKWEVDFPCFCFHLNMAWEEPTDEI VVIGSCMTDPDSIFNECDGLK  
SVLSEIRLNKTKGSTRKAI IENPDEOVLEAGMVRNKLGRKTOYALATAEPPKV  
SGFAKVDLFTGEVEKFLGDNKYGSEPLFLPRDPSKEEDDGYILAFVHDEKEMTSEL  
QIVNATILKEATVKLPSRVPGYGHGTFINANDLANQA"  
BASE COUNT 667 a 437 c 433 g 627 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2,74e-157 Length: 2164  
Score: 2291.00 Matches: 433  
Percent Similarity: 81.97% Conservative: 67  
Best Local Similarity: 70.98% Mismatches: 92  
Query Match: 72.73% Indels: 18  
DB: 8 Gaps: 5  
US-09-758-269-6 (1-599) x STU276244 (1-2164)  
QY 1 MetaLaserPheThrAlaThrAlaAlaValSerGlyArgTyrLeuGlyGlyAsnHisThr 20  
|||  
Db 67 GTAGCTATGGCCACTACTACTTTCACATGCTACAAATACATGATT----- 111  
|||  
QY 21 GlnProProLeuSerSerSerGlnSerSerAspLeuSerTyrCysSer-----Ser 37  
|||  
Db 112 AAGCTAAGTTGTCAATGCCATCATCAAAAGAGTTTGGTTTTCATCAAACTCTATTCT 171  
|||  
QY 38 LeuProMetAlaSerArgValThrArgLysLeuAsnValSerSerAlaLeuHisThrPro 57  
|||  
Db 172 CTACTCAAAAATCAACATAATAGCAAGTCTCAACATTAACTCTCTCTCTCAAGCTCCA 231  
|||  
QY 58 ProAlaLeuHisPheProLysGlnSerSerAsn-----SerProAlaLleVal 73  
|||  
Db 232 CCTATTCTTCATTTTCTTAACAATCTTCAATATTAACAACACCAAACTAGTACAATT 291  
|||  
QY 74 VallysProLysAlaLysGluSerAsnThr-----LysGlnMetAsnLeuPhe 89  
|||  
Db 292 TCACACCCAAAACAAGAAAACAACACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 351  
|||  
QY 90 GlnArgAlaAlaAlaAlaLeuAspAlaAlaGluGlyPheLeuValSerHisGluLys 109  
|||  
Db 352 CAGAAAGCAGCAGCAATGGCTTTAGTAGTCTGTAGAAGGTGCTTTAAACAAAACACCACTC 411  
|||

Asteridae; lamids; Solanales; Solanaceae; Solanum.

1 Burbridge, A., Taylor, I.B. and Thompson, A.

Potato putative 9-cis-epoxycarotenoid dioxygenase 1 cDNA

Unpublished

2 (bases 1 to 2164)

Burbridge, A.

Direct Submission

Submitted (02-MAR-2000) Burbridge A., Plant Science Division, The

University of Nottingham, Sutton Bonington Campus, Loughborough,

Leicestershire, LE12 5RD, UNITED KINGDOM

Location/Qualifiers

1..2164

/organism="Solanum tuberosum"

/mol\_type="mRNA"

/db\_xref="taxon:4113"

1..2164

/gene="nced1"

73..1887

/gene="nced1"

/function="cleavage of 9-cis-epoxycarotenoids"

/codon\_start=1

/product="putative 9-cis-epoxycarotenoid dioxygenase"

/protein\_id="CAB76920.1"

/db\_xref="GI:7209269"

/translation="MATTTSHTNTWIKPLSPFSSKEFGFASNSISLLKQHNQRL  
NINSLOAPPILHPKOSNNYQTKPTSTISHPKQNNSSSSISKNVLQVAAAALD  
AVEGALTGHELEHLPKTPADPRVOISGNFAPVPENPCQSLPTVKIKPCVGVVVRN  
GANLPFETAGRHFDGDMVHAFQKNGSASACRFTETEFVQEKALGRPVFKAI  
GELHSGIARLMLFYARGLFGLDHSRGTVANGLVYFNNRLLMSEDDLLPHVKV  
TPTGDLTGREFDFDQKSTIAHPKLDPSGELFALSVDIOPKYLKRYFRSKNGE  
KSNDEIPEVEDPTMMHDFAI TEKFVII PQQVFMSEMRGGSFVVDYKNSRFGI  
LDKAKDGSDLKWEVDFPCFCFHLNMAWEEPTDEI VVIGSCMTDPDSIFNECDGLK  
SVLSEIRLNKTKGSTRKAI IENPDEOVLEAGMVRNKLGRKTOYALATAEPPKV  
SGFAKVDLFTGEVEKFLGDNKYGSEPLFLPRDPSKEEDDGYILAFVHDEKEMTSEL  
QIVNATILKEATVKLPSRVPGYGHGTFINANDLANQA"

```
Qy 110 LeuHisProLeuProLysThrAlaAspProSerValGlnIleAlaGlyAsnPheAlaPro 129
Db 412 GAGCACCCTTTGCGAANAACAGCCAGCCAGAGTCAGATTTCTGGGAATTTTCGCTCCG 471
Qy 130 ValAsnGluGlnProValArgAsnLeuProValValGlyLysLeuProAspSerIle 149
Db 472 GTACCGGAACACCAGTTTGCCAATCTCTTCGCGTCACCGGAAAAATACCCAAATGTGT 531
Qy 150 LysGlyValTyrValArgAsnGlyAlaAsnProLeuHisGluProValThrGlyHisHis 169
Db 532 CAAGCGGTTCAGTTCGAAACAGGAGCTAACCCCTCTTTTCGAACCAACCGCGGACGCCAT 591
Qy 170 PhePheAspGlyAspGlyMetValHisAlaValLysPheGluHisGlySerAlaSerTyr 189
Db 592 TTCTTCGACGGCAGGTATGGTTACCCCGTTCAATTCAAAATGGCTAGCTAGTTTAC 651
Qy 190 AlaCysArgPheThrGlnThrAsnArgPheValGlnGluArgGlnLeuGlyArgProVal 209
Db 652 GCTTGCGGTTCACTGAAACAGAGAGGTTTGTTCGAAGAAAAAGCTTTGGGTGCGCCTGTT 711
Qy 210 PheProLysAlaIleGlyGluLeuHisGlyHisThrGlyIleAlaArgLeuMetLeuPhe 229
Db 712 TTCCTTAAGCCATTGGTGAATTACATGCTCACTCTGGAATTGCAAGGCTTATGCTGTT 771
Qy 230 TyrAlaArgAlaAlaAlaGlyIleValAspProAlaHisGlyThrGlyValAlaAsnAla 249
Db 772 TACGCTCGGGCTCTCGGACTTATTGATCACAGTAGAGGAAGTGGAGTTGCAACGCC 831
Qy 250 GlyLeuValTyrPheAsnGlyArgLeuLeuAlaMetSerGluAspAspLeuProTyrGln 269
Db 832 GGTTAGTCTATTTCATTAACCGATTACTTCTGATGCTGAAGATGATTTGCTTACCAT 891
Qy 270 ValGlnIleThrProAsnGlyAspLeuLysThrValGlyArgPheAspPheAspGlyGln 289
Db 892 GTAAAGGTAAACACCCCGCGGATCTTAAAAACAGAGGGTCGATTTCGATTCGACGGCCAG 951
Qy 290 LeuGluSerThrMetIleAlaHisProLysValAspProGluSerGlyGluLeuPheAla 309
Db 952 CTAAATCCCATGATAGCTCACCAAGCTCAGCCAGCTTCCGGTGAGCTATTGCT 1011
Qy 310 LeuSerTyrAspValValSerLysProTyrLeuLysTyrPheArgPheSerProAspGly 329
Db 1012 CTAGCTACGATGTGATTTCAGAGCCATACCTCAAGTATTTTCAGATTTTCAAAAATGGG 1071
Qy 330 ThrLysSerProAspValGluIleGlnLeuAspGlnProThrMetMetHisAspPheAla 349
Db 1072 GAAAAATCAAAATGATGTTGAAATCCCATTTGAAGACCCCAACAATGATGATTTCCGCA 1131
Qy 350 IleThrGluAsnPheValValValProAspGlnGlnValValPheLysLeuProGluMet 369
Db 1132 ATTACTGAAAAATTCGTCATCATCTCTGATCAACAAGTCGTTTTCAAGATGCTCGAATG 1191
Qy 370 IleArgGlyGlySerProValValTyrAspLysAsnLysValAlaArgPheGlyIleLeu 389
Db 1192 ATCCGTGGAGGTTCACCGGTGTTTACGACAAGAACAAAGGTGTCCTGATTTGGGATCTG 1251
Qy 390 AspLysTyrAlaGluAspSerSerAsnIleLysTrpIleAspAlaProAspCysPheCys 409
Db 1252 GATAAGTACGCGAAGATGGGTCTGATTGAAATGGGTGGAAGTTCCCTGATTTGTTCTGT 1311
Qy 410 PheHisLeuTrpAsnAlaTrpGluLeuProGluThrAspGluValValIleGlySer 429
Db 1312 TTCATCTCTGGAATGCTTGGGAAGAACCAAGAACAGATGAATGTTGTAATTTGGTTCA 1371
Qy 430 CysMetThrProProAspSerIlePheAsnGluSerAspGluAsnLeuLysSerValLeu 449
Db 1372 TGTATGACACACACAGACTCCATTTCAATGAATGATGATGAGGACTAAAAAGCGTTTA 1431
Qy 450 SerGluIleArgLeuAsnLeuLysThrGlyGluSerThrArgArgProIleIleSerAsn 469
Db 1432 TCCGAAATTCGTCCTCAATTTGAAAAACAGGGAATCAACAAGAAAGCCATTAATCGAAAC 1491
Qy 470 GluAspGlnGlnValAsnLeuGluAlaGlyMetValAsnArgAsnMetLeuGlyArgLys 489
```

```
Db 1492 CCGGATCAACAAGTGAATTTAGAGGCTGGAATGGTGAACAGAAACAACACTTGAAGGAA 1551
Qy 490 ThrLysPheAlaTyrIleuAlaLeuAlaGluProTrpProLysValSerGlyPheAlaLys 509
Db 1552 ACACAGTATGCTTATTGGCTATTGCTGAACACCATGGCCAAAAGTTTCTGTTTTGCAAAA 1611
Qy 510 ValAspLeuThrThrGlyGluValLysLysHisLeuTyrGlyAspAsnArgTyrGlyGly 529
Db 1612 GTAGACCTGTTTACAGGTGAAGTTGAGAAATTCATTATTTGTCACAAATATGGTGGG 1671
Qy 530 GluProLeuPheLeuProGlyGlu-----GlyGlyGluGluAspGlyGlyIleLeu 547
Db 1672 GAGCCTCTTTTACCAAGACCCCAACAGGAGGAGGACGATGGTTATATTTA 1731
Qy 548 CysPheValHisAspGluLysThrTrpLysSerGluLeuGlnIleValAsnAlaValSer 567
Db 1732 GCTTTCGTTTCATGATGAGAAAGATGACATCAGAGCTTCAAAATTTGTTAATGCAATGACT 1791
Qy 568 LeuGluValGluAlaThrValLysLeuProSerArgValProTyrGlyPheHisGlyThr 587
Db 1792 TTAAGTTAGAGGCAACAGTGAAACTTCCATCAAGAGTTCCTTATGTTTTTCATGGAAT 1851
Qy 588 PheIleGlyAlaAspAspLeuAlaLysGln 597
Db 1852 TTCATAATGCAATGATTGGCAATCAG 1881

RESULT 7
LOCUS LENEOKANT 2171 bp mRNA linear PLN 22-MAY-1998
DEFINITION Lycopersicon esculentum mRNA for nine-cis-epoxycarotenoid
dixygenase.
ACCESSION 297215
VERSION 297215.1 GI:2769641
KEYWORDS nine-cis-epoxycarotenoid dioxygenase.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; lamids; Solanales; Solanales; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE
1 Burbridge, A., Grieve, T. M., Jackson, A., Thompson, A. and Taylor, I. B.
Structure and expression of a cDNA encoding a putative neoxanthin
cleavage enzyme (NCE) isolated from a wilt-related tomato
(Lycopersicon esculentum Mill.) library
J. Exp. Bot. 47, 2111-2112 (1997)
2 Burbridge, A.
Direct Submission
Submitted (01-JUL-1997) Burbridge A., The University of Nottingham,
Physiology and Environmental Science, Sutton Bonington Campus,
Loughborough, Leicestershire, LE12 5RD, UK
Revised by [3]
3 (bases 1 to 2171)
Burbridge, A.
Direct Submission
Submitted (09-JAN-1998) Burbridge A., The University of Nottingham,
Physiology and Environmental Science, Sutton Bonington Campus,
Loughborough, Leicestershire, LE12 5RD, UK
On Jan 13, 1998 this sequence version replaced gi:2243153.
COMMENT
FEATURES
Location/Qualifiers
source
1..2171
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
79..1896
/codon_start=1
/product="nine-cis-epoxycarotenoid dioxygenase"
/protein_id="CAB10186.1"
/db_xref="GI:2769642"
/db_xref="SPTREMBL:O24023"
/translation="MATTTSHTATNTWTKLSPSPSKFSGFASNSISLLKNQHNQSL
NINSSLAQPFILHFPKQSSNYQTPKNTTISHPKENNSSSSSTSKWNLVQKAAAMAL"

CDS
```



```
RESULT 8
LES439079      19018 bp      DNA      linear      PLN 13-MAR-2003
LOCUS          Lycopersicon esculentum nced1 gene for 9-cis-epoxycarotenoid
DEFINITION     dioxygenase.
ACCESSION      AJ439079
VERSION        AJ439079.2 GI:28974076
KEYWORDS       9-cis-epoxycarotenoid dioxygenase; nced1 gene.
SOURCE         Lycopersicon esculentum (tomato)
ORGANISM       Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE      1 Thompson,A.J.
AUTHORS        Completionation of notabilis, a tomato mutant deficient in the
TITLE          abscisic acid biosynthetic enzyme 9-cis-epoxycarotenoid dioxygenase
JOURNAL        Unpublished
REFERENCE      2 (bases 1 to 19018)
AUTHORS        Thompson,A.J.
TITLE          Direct Submission
JOURNAL        Submitted (13-MAR-2002) Thompson A.J., Plant Genetics and
REMARK         Biotechnology, Horticulture Research International, Wellesbourne,
COMMENT        Warwick, CV35 9BF, UNITED KINGDOM
FEATURES       revised by author [14-MAY-2002]
Related mRNA entry 297215.
Location/Qualifiers
1..19018
/organism="Lycopersicon esculentum"
/mol_type="genomic DNA"
/cultivar="Money-maker"
/db_xref="taxon:4081"
/chromosome="7"
/map="7q40"
/clone="NCED1.1"
267..567
/rpt_family="tomato anionic peroxidase inverted repeat
(TAPTR)"
repeat_region /rpt_type=INVERTED
2343..3038
/rpt_type=DIRECT
gene           5304..7570
TATA_signal   5304..5311
TATA_signal   5334..5337
mRNA          5404..7570
CDS           5482..7299
/gene="nced1"
/function="abscisic acid biosynthesis"
/codon_start=1
/product="9-cis-epoxycarotenoid dioxygenase"
/protein_id="CAD30202.1"
/db_xref="GI:28974077"
/translatability="MATTSHATNTWTIKLSMPSSKEFGFASNSISLLKNQNRQSL
NINSQAPLILHPKQSSNYQTTPKNNITISHPKQENNSSSSTSKWNLVQKAAAMAL
DAVSAALTKELEHPLKPTADPRVQISGNFAPVPENPVCSLPVTGKIPKCVQGVYVR
NGANPLPEPTAGHHFFDGDGMVHVAQPKNGSASVACRFTETRLVQEKALGRVPFPA
IGELHGSGIARLMLFYARGLFGVDHSKGTGVANAGLVYFNNLLAMSEDDLPLPHVK
VTPGDLKTEGRFEDGGLKSTIAHPKLDPSVSGELFALSVDVTQKPLKYPFRSKNG
EKSNDVEIPSDPTMMHDFALTENFVPIPDQOVFKMSEMRGSPVYIDKNKVRSEGL
ILDYAKDGLDKLWVEVDFCFHLNMAWEAEATDEIVIGSCMTPPDPSIFNECDGL
KVSLEIRLNLTGKSTRKSIENPDEQVLEAGVARNKILGRTEYAYLAIAEPWK
VSGFAKNLFTGEVKEFTYGDNKGGEPLFLPRDPNSKEEDDGYILAFVHDEKEWKE
LQIVNANSKLKLEATVKLPSRVPYGFHGTFFINANDLANQA"
7570
/gene="nced1"
polyA_site    /evidence=experimental
8285..8575
repeat_region /rpt_family="tomato anionic peroxidase inverted repeat
(TAPTR)"
```



```

Qy 249 AlaGlyLeuValTyrPheAsnGlyArgLeuLeuAlaMetSerGluAspLeuProTyr 268
Db 6241 GCCGGTGTAGCTATTTCATTAACCGATTACTTGTCTGTAAGATGATTTGCTTAC 6300
Qy 269 GlnValGlnIleThrProAsnGlyAspLeuLeuValGlyArgPheAspGly 288
Db 6301 CATGTAAGGTAAACACCCCGCGCATCTTAAACAGAGGGTGGATTCGATTTCCAGCGC 6360
Qy 289 GlnLeuGluSerThrMetIleAlaHisProLysValAspProGluSerGlyGluLeuPhe 308
Db 6361 CAGCTAAATCCACCATGATAGCTACCCAAAGCTCGACCCAGTTCCGGTGAGCTATT 6420
Qy 309 AlaLeuSerTyrAspValValSerLysProTyrLeuLysTyrPheArgPheSerProAsp 328
Db 6421 GCTCTTAGCTAGCTAGTATTCAGAAAGCCATACCTCAAGTACTTCAGATTTTCAAAAAAT 6480
Qy 329 GlyThrLysSerProAspValGluIleGlnLeuAspGlnProThrMetMetHisAspPhe 348
Db 6481 GGGGAAAAATCAATGATGTTGAAATTCGATTGAAGACCCCAACAAATGATGATGATTTTC 6540
Qy 349 AlaIleThrGluAsnPheValValProAspGlnGlnValValPheLysLeuProGlu 368
Db 6541 GCATTTACTGAGAACTTCGTCTGCTATTCTGTATCAACAGTGGTTTCAAGATGCTGAA 6600
Qy 369 MetIleArgGlyGlySerProValValTyrAspLysAsnLysValAlaArgPheGlyIle 388
Db 6601 ATGATCCGTGGAGGTTCACCGGTGTTTACGACAAAGCAAAAGTTTCCCGATTTGGTATT 6660
Qy 389 LeuAspLysTyrAlaGluAspSerSerAsnIleLysTyrIleAspAlaProAspCysPhe 408
Db 6661 CTGGATAAGTACGCCAAGATGGGTCTGATTTGAAATGGTTGAAAGTACCTGATTTGTTTC 6720
Qy 409 CysPheHisLeuTrpAsnAlaTrpGluGluProGluThrAspGluValValIleGly 428
Db 6721 TGTTTCCACCTCTGGAATGCTTTGGAGAGCAGAGAAACAGATGAATCGTTGTAATGGT 6780
Qy 429 SerCysMetThrProProAspSerIlePheAsnGluSerAspGluAsnLeuLysSerVal 448
Db 6781 TCATGTATGACACCAACAGATCCATTTTCAATGAATGTGTAAGGGCTAAAGAGTGT 6840
Qy 449 LeuSerGluIleArgLeuAsnLeuLysThrGlyGluSerThrArgArgProIleIleSer 468
Db 6841 TTATCCGAAATCCGCTCAATTTGAAACAGCGGAAATCAACAGAAATTCATTAATCGAA 6900
Qy 469 AsnGluAspGlnGlnValAsnLeuGluAlaGlyMetValAsnArgAsnMetLeuGlyArg 488
Db 6901 AACCCGGATGAACAGTGAATTTAGAGCTGAATGGTGAACCCGAAACAAATCCGGAAGG 6960
Qy 489 LysThrLysPheAlaTyrIleuAlaLeuAlaGluProTrpProLysValSerGlyPheAla 508
Db 6961 AAAACAGAGTATGCTATTGCTATCGTGAACCATGCCCAAAAGTTTCTGTTTTCGA 7020
Qy 509 LysValAspLeuThrThrGlyGluValLysLysHisLeuTyrGlyAspAsnArgTyrGly 528
Db 7021 AAAGTAAACCTGTTCCCGGTGAAGTTGAGAAATTCATTTATGGTGAACAAATATGTT 7080
Qy 529 GlyGluProLeuPheLeuProGlyGlu-----GlyGlyGluGluAspGluGlyTyrIle 546
Db 7081 GGGGAACCTCTTTTTTACCAGAGACCCCAACACAGAGGAGAGAGAGATGTTTATATT 7140
Qy 547 LeuCysPheValHisAspGluLysThrTrpLysSerGluLeuGlnIleValAsnAlaVal 566
Db 7141 TTAGCTTCTGTCACGATGAGAAAGATGGAATTCAGAAATCAGAAATGTTTAAACGCAATG 7200
Qy 567 SerLeuGluValGluAlaThrValLysLeuProSerArgValProTyrGlyPheHisGly 586
Db 7201 AGTTTGAAGTGGAGCGCACTGTGAAGCTTCCATCAAGAGTTCTCTATGATTTTCATGGA 7260
Qy 587 ThrPheIleGlyAlaAspLeuAlaLysGln 597
Db 7261 ACATTCATAAACGCCAATGATTTGGCAATCAG 7293

```

# RESULT 9 AX148316

LOCUS

DEFINITION

AX148316

VERSION

AX148316.1

KEYWORDS

SOURCE

ORGANISM

Lycopodium esculentum (tomato)

Lycopodium esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Asteridae; lamiales; Solanales; Solanaceae; Solanum; Lycopodium.

1

Iuchi, S., Kobayashi, M. and Shinozaki, K.

Transgenic plants carrying neoxanthin cleavage enzyme gene

Patent: EP 1116794-A 15 18-JUL-2001;

Riken (JP)

Location/Qualifiers

1..1818

/organism="Lycopodium esculentum"

/mol\_type="genomic DNA"

/db\_xref="taxon:4081"

1..1818

/note="unnamed protein product"

/codon\_start=1

/protein\_id="CAC41199.1"

/db\_xref="GI:14347204"

/translation="MATTSHATNTWIKTLMSPSSKEFFASNSISLKNQHNROSL

NINSLQAPLILHFKOSSNYOTPKNNTISHPKOENNNSSSSTSKNWLQKAAAMAL

DAVSALTKELEHPLPKTADPRVQISGNFAPVPENPVCSLPVTGKIPKCVQGVYVR

NGALPEPTAGHHFFDGDGWHAVQFKNGSASYACRFTETERLVQEKALGRPVPPKA

IGELHSGTAGRLMLFARGLVGHDHSGTVANAGLVYFNNRLLLAMSDLPYHVK

VTPGCDLKEGDFDQGLKSTMLIAHPKLDPSVGLFALSVDVIQPKYLKFRFSKNG

EKSNVDIPEVDPTMHDFAITENFVVIPOQVVFVKMSEMRGGSPPVVDKMKVSRFG

ILDKYKDGSLKWEVDPDCCFHLNMAWEAETDELIVIGSCWTPDPSIENECDEGL

KSVLSRLNLKTKSTRKSIENPDQVNLNLEAGVRNKNLGRKTEYALAYLAERWPK

VSGFAKNLFTGEVEKFIYGDNKYGGPFLFPRDPNSKBEDDDGYILAFVHDEKEWKSE

LQIYNAMSLKLEATVKLPSRVPYGFHGTFFINANDLANQA"

BASE COUNT 569 a 383 c 369 g 497 t

ORIGIN

Alignment Scores:

Pred. No.: 1.24e-156 Length: 1818

Score: 2280.50 Matches: 435

Percent Similarity: 82.04% Conservative: 63

Best Local Similarity: 71.66% Mismatches: 90

Query Match: 72.40% Indels: 19

DB: 6 Gaps: 6

US-09-758-269-6 (1-599) x AX148316 (1-1818)

Qy 5 ThrAlaThrAlaAlaValSerGlyArgTrpLeuGlyGlyAsnHisThrGlnProLeu 24

Db 7 ACTACTACTTCACATGCCCAAAATACATGGATT-----AAGACTAAGTTG 51

Qy 25 SerSerSerGlnSerSerAspLeuSerTyrCysSer-----SerLeuProMetAla 41

Db 52 TCAATGCCATCATCAAGAGGAGTTGGTTTTCATCAAACTCTATTCTCTACTCAAAAT 111

Qy 42 SerArgValThrArgLysLeuAsnValSerSerAlaLeuHisThrProAlaLeuHis 61

Db 112 CAACATAATAGGCAAGTCTCAACATTAATTTCTCTCTTCAAGCTCCACCTACTTCTAT 171

Qy 62 PheProLysGlnSerSerAsn-----SerProAla-----IleValValLysProLys 77

Db 172 TTTCTTAACCAATCTTCAATATCAACACCAAGAAATAATACATTTTCACCCCAAAA 231

Qy 78 AlalysGluSerAsn-----ThrLysGlnMetAsnLeuPheGlnArgAla 92

Db 232 CAAGAAACAACAACACTCTCTCTTCTTCACTTCCCAAGTGAATTTTAGTCGAGAAGCA 291

Qy 93 AlalaAlaAlaLeuAspAlaAlaGlyPheLeuValSerHisGluLysLeuHisPro 112

```
Db 292 GCAGCAATGGCTTTAGATGCTGTAGAAAGTGCCTTTAACTAAACATGAACCTTGAACACCCCT 351
Qy 113 LeuProLysThrAlaAspProSerValGlnIleAlaGlyAsnPheAlaProValAsnGlu 132
Db 352 TTGCCGAAACAGCCGACCCAGAGTCAGATTTCTGGGAATTTTGTCTCCGGTACCGGAA 411
Qy 133 GlnProValArgAsnLeuProValValGlyLysLeuProAspSerIleLysGlyVal 152
Db 412 AATCAGTCTGTCAATCTCTTCCTCCGGTACCCGGAATAATACCCTAAATGTTGTTAGGGCGTT 471
Qy 153 TyrValArgAsnGlyAlaAsnProLeuHisGluProValThrGlyHisHisPhePheAsp 172
Db 472 TACGTTCCGAACCGAGAGCTAACCTCTTTTGAACCAACCCGCGACACCACTTTCTTCGAC 531
Qy 173 GlyAspGlyMetValHisAlaValLysPheGluHisGlySerAlaSerTyrAlaCysArg 192
Db 532 GCGACCGGTATGTTTACCGCTTCAATTCACCAAAATGGGTCCGCTAGTTTACGCTTCCGCT 591
Qy 193 PheThrGlnThrAsnArgPheValGlnGluArgGlnLeuGlyArgProValPheProLys 212
Db 592 TTCACTGAACAGAGAGGCTTGTTCAGAAAAGAGCTTTGGGTCCGCTTTTCCCTAAA 651
Qy 213 AlaIleGlyGluLeuHisGlyHisThrGlyIleAlaArgLeuMetLeuPheTyrAlaArg 232
Db 652 GCCATTGGTGAATTACATGGTCACTCTGGAATTGCAAGGCTTATGCTGTTTACGCTCGT 711
Qy 233 AlaAlaAlaGlyIleValAspProAlaHisGlyThrGlyValAlaAsnAlaGlyLeuVal 252
Db 712 GGGCTCTTCGGACTTGTGTATCAGTAAGAAAGAACTGGTGTGCAAAACGCCGGTTTAGTC 771
Qy 253 TyrPheAsnGlyArgLeuLeuAlaMetSerGluAspAspLeuProTyrGlnValGlnIle 272
Db 772 TATTTTCAATAACCGATTACTTGTCTATGCTGCTGAAAGATGATTTTGGCTTACCATGTAAAGGTA 831
Qy 273 ThrProAsnGlyAspLeuLysThrValGlyArgPheAspPheAspGlyGlnLeuGluSer 292
Db 832 ACACCCACCGCGGATCTTTAAACAGAGGCTCGATTTCGATTTTCGACGGCCAGCTAAATCC 891
Qy 293 ThrMetIleAlaHisProLysValAspProGluSerGlyGluLeuPheAlaLeuSerTyr 312
Db 892 ACCATGATAGTCAACCAAGCTCGACCACTGTCGGTGAAGTATTTTGTCTTCTTAGCTAC 951
Qy 313 AspValValSerLysProTyrLeuLysTyrPheArgPheSerProAspGlyThrLysSer 332
Db 952 GATGTGATTCAGACCCATACCTCAAGTACTTCAGATTTTCAAAATATGGGAAAAATCA 1011
Qy 333 ProAspValGluIleGlnLeuAspGlnProThrMetHisAspPheAlaIleThrGlu 352
Db 1012 AATGATGTTGAAATTCAGATTGAAGACCAACAATGATGATGATTTTCGCAATTTACTGAG 1071
Qy 353 AsnPheValValValProAspGlnGlnValValPheLysLeuProGluMetIleArgGly 372
Db 1072 AACTTCGTCGTCATTCTCTGATCAACAAGTCGTTTTTCAAGATGTCTCGAAATGATCCGTGGA 1131
Qy 373 GlySerProValValTyrAspLysAsnLysValAlaArgPheGlyIleLeuAspLysTyr 392
Db 1132 GGTTCACCGGTGGTTTACGACAAACAAGTTTCCCGATTTGGTATTCTGGATAAGTAC 1191
Qy 393 AlaGluAspSerSerAsnIleYrIleAspAlaProAspCysPheCysPheHisLeu 412
Db 1192 GCGAAAGATGGTCTGAATTTGAAATGGGTTCAAGTACCTGATGTTTCTGTTTCCACCTC 1251
Qy 413 TrpAsnAlaTrpGluProGluThrAspGluValValIleGlySerCysMetThr 432
Db 1252 TGGAAATGCTTGGGAAGACGACAGAAACAGATGAAATCGTTGTAATTTGGTTCATGTATGACA 1311
Qy 433 ProProAspSerIlePheAsnGluSerAspGluAsnLeuLysSerValLeuSerGluIle 452
Db 1312 CCACAGACTCCATTTTCAATGATGTGATGAAGGGCTAAAGAGTGTTTTATCCGAATC 1371
Qy 453 ArgLeuAsnLeuLysThrGlyGluSerThrArgArgProIleIleSerAsnGluAspGln 472
Db 1372 CGTCTCAATTTGAAAACAGGGAATCAACAAGAAAAATCCATAATCGAAAAACCCCGGATGAA 1431
```

```
Qy 473 GlnValAsnLeuGluAlaGlyMetValAsnArgAsnMetLeuGlyArgLysThrLysPhe 492
Db 1432 CAAGTGAATTTAGAAAGCTGGAATGGTGAACCGAAACAAACTTCGGAAGAAACAGAGTAT 1491
Qy 493 AlaTyrLeuAlaLeuAlaGluProTyrProLysValSerGlyPheAlaLysValAspLeu 512
Db 1492 GCTTATTTGGCTATCGCTGCAACCATGGCCAAAGATTCTCGTTTGGAAAAGTAAACCTG 1551
Qy 513 ThrThrGlyGluValLysLysHisLeuTyrGlyAspAsnArgTyrGlyGlyGluProLeu 532
Db 1552 TTCACCGGTCAAGTTGAGAAATTCATTATATGGTGACAAACAATATATGTTGGGAACCTCTT 1611
Qy 533 PheLeuProGlyGlu-----GlyGlyGluGluAspGluGlyTyrIleLeuCysPheVal 550
Db 1612 TTTTACCAAGAGAGACCCCAACGACGAAAGAACACATGCTTATATATTTTACGCTTTCGTT 1671
Qy 551 HisAspGluLysThrTyrLysSerGluLeuGlnIleValAsnAlaValSerLeuGluVal 570
Db 1672 CACGATGAGAAAGATGGAATCAGAACTGCAAAATGTTTAAACGCAATGAGTTTGAAGTTG 1731
Qy 571 GluAlaThrValLysLeuProSerArgValProTyrGlyPheHisGlyThrPheIleGly 590
Db 1732 GAGCAACTGTGAAGCTTCCATCAAGAGTTCCTTATGCAATTCATGGAACATTCATAAAC 1791
Qy 591 AlaAspAspLeuAlaLysGln 597
Db 1792 GCCAATGATTGGCAATCAG 1812

RESULT 10
BD017436
LOCUS BD017436 1818 bp DNA linear PAT 27-AUG-2002
DEFINITION Transgenic plant using neozanthine cleaving enzyme gene.
ACCESSION BD017436
VERSION BD017436.1 GI:22558612
KEYWORDS JP 2001258579-A/8.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum

REFERENCE
1 (bases 1 to 1818)
Iuchi, K., Kobayashi, M. and Shinozaki, K.
Transgenic plant using neozanthine cleaving enzyme gene
Patent: JP 2001258579-A 8 25-SEP-2001;
THE INSTITUTE OF PHYSICAL AND CHEMICAL RESEARCH
OS Lycopersicon esculentum (tomato)
PN JP 2001258579-A/8
PD 25-SEP-2001
PF 11-JAN-2001 JP 2001003476
PI KIYOSHI IUCHI, MASATOMO KOBAYASHI, KAZUO SHINOZAKI PC
C12N15/09,A01H5/00,C12N5/10,C12N5/02,C12N15/00,C12N5/00 CC
Transgenic plant using neozanthine cleaving enzyme gene FH Key
COMMENT OS Lycopersicon esculentum (tomato)
Location/Qualifiers
(1) . (1818).
FEATURES
source
Location/Qualifiers
1. .1818
/organism="Lycopersicon esculentum"
/mol_type="genomic DNA"
/db_xref="taxon:4081"
BASE COUNT 569 a 383 c 369 g 497 t
ORIGIN
Alignment Scores:
Pred. No.: 1.24e-156 Length: 1818
Score: 2280.50 Matches: 435
Percent Similarity: 82.04% Conservative: 63
Best Local Similarity: 71.66% Mismatches: 90
Query Match: 72.40% Indels: 19
DB: Gaps: 6
US-09-758-269-6 (1-599) x BD017436 (1-1818)
```



Qy	5	ThrAlaThrAlaAlaValSerGlyArgTrpLeuGlyGlyAsnHisThrGlnProProLeu	24
Db	7	ACTACTACTTCACATCGCCAAATACATACGAGTT	51
Qy	25	SerSerSerGlnSerSerAspLeuSerTyrCysSer	41
Db	52	TCAATGCCATCATCAAGAGAGTTGGTTTGGCATCAAACTCTATTCTCTACTCAAAAT	111
Qy	42	SerArgValThrArgLysLeuAsnValSerSerAlaLeuHisThrProProAlaLeuHis	61
Db	112	CAACATAATAGCAAGAGCTCAACATTAATTTCTCTCTTCAAGCTCCACCTATATCTTCAT	171
Qy	62	PheProLysGlnSerSerAsn	77
Db	172	TTTCTTAACAATCTCTCAATTTATCAACACCAAGAAATAACAATTTCAACCCCAAAA	231
Qy	78	AlaLysGluSerAsn	92
Db	232	CAAGAAACAACAACCTCTCTTCTTCAACTTCCAAAGTGGAAATTTAGTCGCAAGAGCA	291
Qy	93	AlaAlaAlaLeuAspAlaAlaGluGlyPheLeuValSerHisGluLysLeuHisPro	112
Db	292	GCAGCAATGGCTTTAGATGCTGTAGAAAGTCTTTAACTAAACATGAACCTTGAACACCT	351
Qy	113	LeuProLysThrAlaAspProSerValGlnIleAlaGlyAsnPheAlaProValAsnGlu	132
Db	352	TTGCGCAAAACAGCGACCCACGAGTCCAGATTTCTGGCAATTTTGTCTCGGTACCGAA	411
Qy	133	GlnProValArgArgAsnLeuProValValGlyLysLeuProAspSerIleLysGlyVal	152
Db	412	AATCCAGTCTGTCAATCTTCTCCGGTCAACCGGAAATAATACCAAAATGTGTTCAAGGGCTT	471
Qy	153	TyrValArgAsnGlyAlaAsnProLeuHisGluProValThrGlyLysHisPhePheAsp	172
Db	472	TACGTTCGAAACGGAGCTAACCTCTTTTGTGAACCAACCGCGGACACCATTTCTTCGAC	531
Qy	173	GlyAspGlyMetValHisAlaValLysPheGluHisGlySerAlaSerTyrAlaCysArg	192
Db	532	GGCAGCGGTATGGTTTACCGCGTTCAATTTCAAAAAATGGTTCGGCTAGTTACGCTTGGCGT	591
Qy	193	PheThrGlnThrAsnArgPheValGlnGluArgGlnLeuGlyArgProValPheProLys	212
Db	592	TTTCTGAAACAGAGAGCTTTGTTCAAGAAAAAGCTTTGGGTGCGCTGTTTTCCTTAAA	651
Qy	213	AlaIleGlyGluLeuHisGlyHisThrGlyIleAlaArgLeuMetLeuPheTyrAlaArg	232
Db	652	GCCATTGGTGAATTACATGCTCACTTGGAAATTCGAAGCTTATGCTGTTTACGCTCGT	711
Qy	233	AlaAlaAlaGlyIleValAspProAlaHisGlyThrGlyValAlaAsnAlaGlyLeuVal	252
Db	712	GGGCTCTTCGGACTTGTGTGATCATCAAGAAAGAACTGGTGTGCAAAACGCGGTTTAGTC	771
Qy	253	TyrPheAsnGlyArgLeuLeuAlaMetSerCluAspAspLeuProTyrGlnValGlnIle	272
Db	772	TATTTCAATAACCCGATTACTTGTCTATGTCTGAAGATGATTTGGCTTACCATGTAAAGGTA	831
Qy	273	ThrProAsnGlyAspLeuLysThrValGlyIleArgPheAspPheAspGlyGlnLeuGluSer	292
Db	832	ACACCCACCGCGATCTTAAACAGAGGCTCGATTCGATTTTCGACGCCAGCTTAAATTC	891
Qy	293	ThrMetIleAlaHisProLysValAspProGluSerGlyGluLeuPheAlaLeuSerTyr	312
Db	892	ACCATGATAGTCAACCCAAAGCTCGACCCAGTTTCCGGTGAGCTATTGCTCTTAGCTAC	951
Qy	313	AspValValSerLysProTyrLeuLysTyrPheArgPheSerProAspGlyThrLysSer	332
Db	952	GATGTGATTCAAGAGCCATACCTCAAGTACTTTCAGATTTTCAAAAAATGGGAAAAATCA	1011
Qy	333	ProAspValGluIleGlnLeuAspGlnProThrMetHisAspPheAlaIleThrGlu	352
Db	1012	AATGATGTTGAAATTCAGTTGAAGACCCCAACAATGATGATGATTTCCGAATTTACTTGAG	1071
Qy	353	AsnPheValValValProAspGlnGlnValValPheLysLeuProGluMetIleArgGly	372

Qy	114	ProLysThrAlaAspProSerValGlnIleAlaGlyAsnPheAlaProValAsnGluGln	133
Db	520	CCCCAAACGGCGGACCCCGAAAGTCCAATCGCCGGAACTTCGCGCGGTGCCGGACAC	579
Qy	134	ProValArgArgAsnLeuProValValGlyLysLeuProAspSerIleLysGlyValTyr	153
Db	580	GCCGCGGATCAAGCCCTTCGGTGGTTCGGAATAATCCCAATGCATTGACGGGTGTAC	639
Qy	154	ValArgAsnGlyAlaAsnProLeuHisGlyProValThrGlyHisHisPhePheAspGly	173
Db	640	GTGCGCAACGGCGGAATCCGCTCTACGAGCCGTCGCGGGCACCACTTCTTCGACGCG	699
Qy	174	AspGlyMetValHisAlaValLysPheGluHisGlySerAlaSerTyrAlaCysArgPhe	193
Db	700	GACGGCATGTCACAGCCGCGAAATTACGAAACGGCGTCGACGTCAGCGTGCCTTC	759
Qy	194	ThrGlnThrAsnArgPheValGlnGluArgGlnLeuGlyArgProValPheProLysAla	213
Db	760	ACCGAAACACAGCGGTCTCGCGCAAGAAATCTTTAGCGCCCGCGGTTCCTCCGAAGGCT	819
Qy	214	IleGlyGluLeuHisGlyHisThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAla	233
Db	820	ATCGCGGAATCTCACGGCCACTCGGCGATAGCTGCCTCTCTCTTTCTACGCTCGCAGC	879
Qy	234	AlaAlaGlyIleValAspProAlaHisGlyThrGlyValAlaAsnAlaGlyLeuValTyr	253
Db	880	CTCTTCCAGCTCGTTGATGGTCCACCGCATGGGGTGGCGAAGCCGGTCTCGTCTAT	939
Qy	254	PheAsnGlyArgLeuLeuAlaMetSerGluAspAspLeuProTyrGlnValGlnIleThr	273
Db	940	TTCAACAAACACACTCTTTGGCCATGTCCGAGACGATTTACCATACCACTGAGAAATCACC	999
Qy	274	ProAsnGlyAspLeuLysThrValGlyArgPheAspPheAspGlyGlnLeuGluSerThr	293
Db	1000	TCTAACGGCGCACTTAACACCGTTGGCGGTGTACGACTTCAACGGCCAGCTCACTCCACA	1059
Qy	294	MetIleAlaHisProLysValAspProGluSerGlyGluLeuPheAlaLeuSerTyrAsp	313
Db	1060	ATGATCGCCACCCGAAACTCGACCCCGTCAACGGCGACCTCCACGGCTCAGCTACGAC	1119
Qy	314	ValValSerLysProTyrLeuLysTyrPheArgPheSerProAspGlyThrLysSerPro	333
Db	1120	GTCTTCAGAAAGCCTTACCTCAAGTACTTCGTTTCTCCGCTGACGGCGTAAAGTCCCCC	1179
Qy	334	AspValGluIleGlnLeuAspGlnProThrMetHisAspPheAlaIleThrGluAsn	353
Db	1180	GACGTCAATCCCCCTGAAAGACCAACCATGATGCACGATTTTCGCATTAACGGAGAT	1239
Qy	354	PheValValValProAspGlnGlnValValPheLysLeuProGluMetIleArgGlyGly	373
Db	1240	TTCTGCTGTCCTCCCGACCAAGCAGTGGTCTTCAAACTAACAGAGATGATCACCGCGCGC	1299
Qy	374	SerProValValTyrAspLysAsnLysValAlaArgPheGlyIleLeuAspLysTyrAla	393
Db	1300	TCCCCCTGCTACTACGACAAAGAAACAAACCTCACGGTTTGGGATTTGGATTAAGAATCGC	1359
Qy	394	GluAspSerSerAsnIleLysTrpIleAspAlaProAspCysPheCysPheHisLeuTrp	413
Db	1360	AAGCAGCGAATGCGATGGGTGGATGCAGCGCCGGAGTGTCTTGCTTCCATCTCTGG	1419
Qy	414	AsnAlaTrpGluGluProGluThrAspGluValValIleGlySerCysMetThrPro	433
Db	1420	AACTGTTGGGAGGACCTGAAACCGACGAGATTGTGGTATTGGTCTCTGCATGACCCCT	1479
Qy	434	ProAspSerIlePheAsnGluSerAspGluAsnLeuLysSerValLeuSerGluIleArg	453
Db	1480	GCGGACTCCATTTTCAATGAATGCGATGAGATTGAAAGACGCTCTGTCTGTGAGATAAG	1539
Qy	454	LeuAsnLeuLysThrGlyGluSerThrArgArgProIleIleSerAsnGluAspGlnGln	473
Db	1540	CTGAACCTTGAGGACGGGCAAGTCACGTCGCGCCCAATTAATCTCC--GACCGCGAACAA	1599
Qy	474	ValAsnLeuGluAlaGlyMetValAsnArgAsnMetLeuGlyArgLysThrLysPheAla	493

Qy	134	ProValArgArgAsnLeuProValValGlyHisGlyLeuProAspSerIleLysGlyValTyr	151
Db	580	GC CGCGGATCAAGCCCTTCGGGTGCGGAAAAATCCCAAAATGCATTGACGGGTGTAC	639
Qy	154	ValArgAsnGlyAlaAsnProLeuHisGlyLeuProValThrGlyHisHisPhePheAspGly	173
Db	640	GTGCGCAACGCGCAATCCGCTCTACGAGCCGCTCGCGGGCACCACTTCTTCGACGCG	699
Qy	174	AspGlyMetValHisAlaValLysPheGluHisGlySerAlaSerTyrAlaCysArgPhe	193
Db	700	GACGGCATGTCACACGCGGTGAAGTTTCAGAAACGCGCTGCACGTACGGGTGCGGCTTC	759
Qy	194	ThrGlnThrAsnArgPheValGlnGluArgGlnLeuGlyArgProValPheProLysAla	213
Db	760	ACCGCAACACAGCGTCTCGGCGCAAGAAATCTTTAGCGCCCGCGGTTCCTCCGAGGCT	819
Qy	214	IleGlyGluLeuHisGlyHisThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAla	233
Db	820	ATCGCGAACTTCACGSCCACTCCGGCATAGCTCGCTCTCTCTCTTCTACGCTCGCAGC	879
Qy	234	AlaAlaGlyIleValAspProAlaHisGlyThrGlyValAlaAlaAsnAlaGlyLeuValTyr	253
Db	880	CTCTTCAGAGTCTGTTGATGGGTCCACCGCATGGGGGTGGCGAAACGCGGTCTCGTCTAT	939
Qy	254	PheAsnGlyArgLeuLeuAlaMetSerGluAspAspLeuProTyrGlnValGlnIleThr	273
Db	940	TTCAACAACCACTCTCTTGGCCATGTCGAGACGATTTACCATACCACGTGAGAAATCACC	999
Qy	274	ProAsnGlyAspLeuLysThrValGlyArgPheAspPheAspGlyGlnLeuGluSerThr	293
Db	1000	TCTAAACGGCGACTTAACACCGTTGGCCGTTACGACTTCAACGCCAGCTCAACTCCACA	1059
Qy	294	MetIleAlaHisProLysValAspProGluSerGlyGluIlePheAlaLeuSerTyrAsp	313
Db	1060	ATGATGCCCCACCGCAAACTCGACCCCGTCAACGGCGACCTTCCACGGCTCAGCTACGAC	1119
Qy	314	ValValSerLysProTyrLeuLysTyrPheArgPheSerProAspGlyThrLysSerPro	333
Db	1120	GTGCTTCAGAAGCCTTACCTCAAGTACTTCGTTTCTCCGCTGACGGCGTAAGTCCCCC	1179
Qy	334	AspValGluIleGlnLeuAspGlnProThrMetHisAspPheAlaIleThrGluAsn	353
Db	1180	GAGCTCGAAATCCCCCTGAAAGACCAACCATGATGACACGATTTCGCAATTAACCGAGAT	1239
Qy	354	PheValValProAspGlnGlnValValPheLysLeuProGluMetIleArgGlyGly	373
Db	1240	TTGCTGCTGTCCCCGACACGAGTGGTCTTCAAACTTAACAGAGATGATCACCGCGCGC	1299
Qy	374	SerProValValTyrAspLysAsnLysValAlaAlaArgPheGlyIleLeuAspLysTyrAla	393
Db	1300	TCCCGCTGCTACGACAAGAACAAACCTCACGGTTTGGATTGGATAAAGAAATGCG	1359
Qy	394	GluAspSerSerAsnIleLysTyrIleAspAlaProAspCysPheCysPheHisLeuTrp	413
Db	1360	AAGACCGCAATGCGGTGCGGTGGATCGACGCGCGGAGTGTCTTCGCTCCCATCTCTGG	1419
Qy	414	AsnAlaTrpGluGluProGluThrAspGluValValIleGlySerCysMetThrPro	433
Db	1420	AACCGTGGGAGGACCTGAAACCCACGAGATTTGTTGATTTGGTCTCTGATGACCCCT	1479
Qy	434	ProAspSerIlePheAsnGluSerAspGluAsnLeuLysSerValLeuSerGluIleArg	453
Db	1480	GCGGACTCCATTTTCAATGAATGCATGAGATTTGAAGACGCTTGTCTCTGAGATAAGG	1539
Qy	454	LeuAsnLeuLysThrGlyGluSerThrArgArgProIleIleSerAsnGluAspGlnGln	473
Db	1540	CTGAACCTTGAGGACGGGCAAGTCAGCTCGCGGCCCACTATATCTCC--GACCGCGAACAA	1599
Qy	474	ValAsnLeuGluAlaGlyMetValAsnArgAsnMetLeuGlyArgLysThrLysPheAla	493

294	MetIleAlaHisProLysValAspProGluSerGlyGluLeuPheAlaLeuSerTyrAsp	313
1060	ATGATGCCACCGGAAACTCGACCCCGTCAACCGCGACCTCCACCGCGCTCAGGTACGAC	1119
314	ValValSerLysProTyrLeuLysTyrPheArgPheSerProAspGlyThrLysSerPro	333
1120	GTGGTTCAGAAAGCTTAACTCAAGTACTTCGTTTCTCCGCTGACGCGGTAAAGTCCCCC	1179
334	AspValGluIleGlnLeuAspGlnProThrMetMethHisAspPheAlaIleThrGluAsn	353
1180	GAGTCTCAAAATCCCCCTGAAAGACCAACCATGATGCACGATTTCGCAATTAACGAGAAAT	1233
354	PheValValValProAspGlnGlnValValPheLysLeuProGluMetIleArgGlyGly	373
1240	TTGCTGCTGTCGCCCGACAGCAGTGGTCTTCAAACTTAACAGAGATGATCACCGGCCGCG	1299
374	SerProValValTyrAspLysAsnLysValAlaAlaArgPheGlyIleLeuAspLysTyrAla	393
1300	TCCCCCGTGGTCTACGACAAGAACCAAACTCACGGTTTGGGATTTGGATAAAGAATGCG	1359
394	GluAspSerSerAsnIleLysTyrIleAspAlaProAspCysPheCysPheHisLeuTrp	413
1360	AAGACGCGAATGCGATGCGGTGGATCGACGCGCGGAGTGTTCCTGCTCCATCTCTCG	1419
414	AsnAlaTrpGluGluProGluThrAspGluValValIleGlySerCysMetThrPro	433
1420	AACCGTGGGAGGACCTGAACCCGACGAGATTGTGTGATTGGGTCTGTGATGACCCT	1479
434	ProAspSerIlePheAsnGluSerAspGluAsnLeuLysSerValLeuSerGluIleArg	453
1480	CGCGACTCCATTTTCAATGAATGCGATGAGAGTTTGAAGACGCTCTGTCTGAGATAAGG	1539
454	LeuAsnLeuLysThrGlyGluSerThrArgArgProIleIleSerAsnGluAspGlnGln	473
1540	CTGAACCTTGAGGACGGGCAAGTCGACTCGCGGCCCATTAATCTCC--GACCGCGAACAA	1599
474	ValAsnLeuGluAlaGlyMetValAsnArgAsnMetLeuGlyArgLysThrLysPheAla	493

334	Qy	ASPVALGILLEGUNDEASPGRINPORNWETHIRASPHALALILRIGIUNSH	335	Qy	ASPVALGILLEGUNDEASPGRINPORNWETHIRASPHALALILRIGIUNSH
1180	Db	GAGTCGAATCCCCTGAAGAGCCAAACCATGATGCACGATTTCGCAATTAACGGAGAA	1235	Qy	ASPVALGILLEGUNDEASPGRINPORNWETHIRASPHALALILRIGIUNSH
354	Qy	PheValValProAspGlnValValPheLysLeuProGluMetIleArgGlyGly	373	Qy	ASPVALGILLEGUNDEASPGRINPORNWETHIRASPHALALILRIGIUNSH
1240	Db	TTGTCGTCTGTCCTCCGACACAGAGTGGTCTTCAACATAACAGAGATGATCACCGCGCGC	1295	Qy	ASPVALGILLEGUNDEASPGRINPORNWETHIRASPHALALILRIGIUNSH
374	Qy	SerProValValTyrAspLysAsnLysValAlaArgPheGlyIleLeuAspLysTyrAla	393	Qy	ASPVALGILLEGUNDEASPGRINPORNWETHIRASPHALALILRIGIUNSH
1300	Db	TCCCCCGTGTCTACGACACAGAACAAACCTCACGGTTTGGGATTTGGATAAGNATCGG	1355	Qy	ASPVALGILLEGUNDEASPGRINPORNWETHIRASPHALALILRIGIUNSH
394	Qy	GluAspSerSerAsnIleLysTyrIleAspAlaProAspCysPheCysPheHisLeuTrp	413	Qy	ASPVALGILLEGUNDEASPGRINPORNWETHIRASPHALALILRIGIUNSH
1360	Db	AAGACCGCGAATGCGATGCGGTGGATCGACGCGCGGAGTGTCTTCCTCCATCTCTGG	1415	Qy	ASPVALGILLEGUNDEASPGRINPORNWETHIRASPHALALILRIGIUNSH
414	Qy	AsnAlaTrpGluGluProGluThrAspGluValValValIleGlySerCysMetThrPro	433	Qy	ASPVALGILLEGUNDEASPGRINPORNWETHIRASPHALALILRIGIUNSH
1420	Db	AACCGTGGGAGGAGCCTGAACCCACGAGATTGTGGTGAATGGGTCTCGCATGACCCCT	1475	Qy	ASPVALGILLEGUNDEASPGRINPORNWETHIRASPHALALILRIGIUNSH
434	Qy	ProAspSerIlePheAsnGluSerAspGluAsnLeuLysSerValLeuSerGluIleArg	453	Qy	ASPVALGILLEGUNDEASPGRINPORNWETHIRASPHALALILRIGIUNSH
1480	Db	GCGGACTCCATTTTCAATGAATGCGATGAGAGATTGAAGACGCTCTGTCTGAGATAAGG	1535	Qy	ASPVALGILLEGUNDEASPGRINPORNWETHIRASPHALALILRIGIUNSH
454	Qy	LeuAsnLeuLysThrGlyGluSerThrArgArgProIleIleSerAsnGluAspGlnGln	473	Qy	ASPVALGILLEGUNDEASPGRINPORNWETHIRASPHALALILRIGIUNSH
1540	Db	CTGAACCTTGAGGACGGGCAAGTCGACTCGCCGCCCATTTATCTCC---GACGCCGAACAA	1595	Qy	ASPVALGILLEGUNDEASPGRINPORNWETHIRASPHALALILRIGIUNSH
474	Qy	ValAsnLeuGluAlaGlyMetValAsnArgAsnMetLeuGlyArgLysThrLysPheAla	493	Qy	ASPVALGILLEGUNDEASPGRINPORNWETHIRASPHALALILRIGIUNSH

```
Db 1597 GTGAACCTTTGAAGCCGCGCATGGTGAACAGAAACAAGCTAGGAGGAAACACTCAGTTCGCT 1656
Qy 494 TyrLeuAlaLeuAlaGluProTyrProLysValSerGlyPheAlaLysValAspLeuThr 513
Db 1657 TATCTGGCTCTCGCGAGCCCTGGCCCAAGTCTCGGGCTTTGCGAAGTGTATTTGTTTC 1716
Qy 514 ThrGlyGluValLysLysHisLeuTyrGlyAspAsnArgTyrGlyGlyProLeuPhe 533
Db 1717 AGTGGGGAAGTGCAGAAAGTACATGATGAGAGAAGAAAGATTCGGTGGGAGCCTCTGTTT 1776
Qy 534 LeuProGlyGluGlyGlyGluGluAspGluGlyTyrLeuLeuCysPheValHisAspGlu 553
Db 1777 CTTCCTCC---AACGGCGAAGAAAGGTGATGGGTATATTTCTGGCATTCGTGCACACGAG 1833
Qy 554 LysThrTrpLysSerGluLeuGlnLeuValAsnAlaValSerLeuGluValGluAlaThr 573
Db 1834 AAAGAGTGGAAATCCGAGCTGCAGATTTGTAATGCCCAAAATTTAAAGCTCGAAGCTTCC 1893
Qy 574 ValLysLeuProSerArgValProTyrGlyPheHisGlyThrPheIleGlyAlaAsp 593
Db 1894 ATCAAACTCCCTCTCGCGTTCCTTATGGTTTTCATGGCATTTCATTCATTCTAAGGAT 1953
Qy 594 LeuAlaLysGln 597
Db 1954 TTGAGGAAGCAA 1965

RESULT 12
AX652128 95769 bp DNA linear PAT 22-MAR-2003
LOCUS AX652128
DEFINITION Sequence 1979 from Patent WO03000898.
ACCESSION AX652128
VERSION AX652128.1 GI:29154942
KEYWORDS
SOURCE Arabidopsis thaliana (chale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
AUTHORS Karagiri, F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, F. and Zou, G.
TITLE Plant genes involved in defense against pathogens
JOURNAL Patent: WO 03000898-A 1979 03-JAN-2003;
SYNGENTA Participations AG (CH)
FEATURES
source 1. 95769
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
BASE COUNT 30958 a 16961 c 16682 g 31168 t
ORIGIN

Alignment Scores:
Pred. No.: 3,456-146 Length: 95769
Score: 2170.50 Matches: 419
Percent Similarity: 80.61% Conservative: 80
Best Local Similarity: 67.69% Mismatches: 87
Query Match: 68.90% Indels: 33
DB: 6 Gaps: 9

US-09-758-269-6 (1-599) x AX652128 (1-95769)
Qy 1 MetAlaSerPheThrAlaThrAlaAlaValSerGlyArgTrpLeu----- 15
Db 33716 ATGGCTTCTACTACTTACTTCTCTTCTACTTCAACCCAAATTCCTCGACCGAATTTCTCA 33775
Qy 16 ---GlyGlyAsnHisThrGlnProLeuProLeuSerSerGlnSerAspLeuSerTyr 34
Db 33776 ACTTCATCTCTCTCTCGAGACCTAACTTCAATCT-----CTATCTTTT 33820
Qy 35 CysSerSerLeuPro-----MetAlaSerArgValThr 45
Db
```

```
Db 33821 TCCTCAACACATTAGAAACAAGAAACTAGTAGTCCCTCTCTATGTTCCTTCCTCCGTCAAC 33880
Qy 46 ArgLysLeuAsnValSerSerAlaLeuHisThr-----ProProAlaLeuHis 61
Db 33881 AAAAAATCTTCGGTCTCTCTCTCTCTCAATCTCCGACATTCAAAACCCACCGTCTTGG--- 33937
Qy 62 PheProLysGlnSerSerAsnSerProAlaIleValValLysProLysAlaLysGluSer 81
Db 33938 -----AAGAAACTATGTAAAGCATGTCAACCACTTGATC---CCCAAAACGACGAATCAG 33988
Qy 82 AsnThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaLeuAspAlaGlu 101
Db 33989 AATCAAAA-----TAAACCCGGTGCAGAGAACCCGCGCTATGGTTTTAGACCGGTGAA 34045
Qy 102 GlyPheLeuValSerHisGlu---LysLeuHisProLeuProLysThrAlaAspProSer 120
Db 34046 AACGCCATGATCTCACAGAGCGCGCGCTCATCCCATCTTAAACCGCGGATCTCTGCC 34105
Qy 121 ValGlnIleAlaGlyAsnPheAlaProValAsnGluGlnProValArgAsnLeuPro 140
Db 34106 GTCCAAATAGCGGGAACCTTCTCCGGTACCGGAGAAACCTGCTCGTGCATAACCTTCG 34165
Qy 141 ValValGlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyAlaAsnPro 160
Db 34166 GTGACTGGAAACAGTACCAGAAATGCATTCAAGGAGTTTACGTACAGAAACGGAGCAATCCA 34225
Qy 161 LeuHisGluProValThrGlyHisPhePheAspGlyAspGlyMetValHisAlaVal 180
Db 34226 CTTCAAAACAGTCTCCGGCCACCATTTGTTTCGACGGAGAGCGGTATGGTACACGCTGTC 34285
Qy 181 LysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheVal 200
Db 34286 CGGTTTCGATAACGGCTCGGTTAGTACGCTTGTGCTTACCGAAACAAACCGGTGGTT 34345
Qy 201 GlnGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGlyHis 220
Db 34346 CAAGAACGTGAATGTGTCGTCGCTTTCCTCCAAAGCAATTTGGAGAGCTTCACGACAT 34405
Qy 221 ThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaAlaGlyIleValAspPro 240
Db 34406 TTAGAAATCGCAAGCTTATGCTCTTCAATACCCCGGGGCTATTTGGGTTAGTTCGACCCG 34465
Qy 241 AlaHisGlyThrGlyValAlaAsnAlaGlyLeuValTyrPheAsnGlyArgLeuAla 260
Db 34466 ACCGAGGACTCGGTGTCGCTAACCGCGTTTATTTATTTCAATGGTCACTCTCTAGCC 34525
Qy 261 MetSerGluAspAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysThr 280
Db 34526 ATGTCGGAAGACGATTTACCGTACCATGTCAAAAGTTACTCAAAACCGGAGATTTAGAAACT 34585
Qy 281 ValGlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysVal 300
Db 34586 TCGSGTTCGTACGATTTTCGACGCTCAGTTAAATCAACATGATAGCCACCCCGAAATC 34645
Qy 301 AspProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValSerLysProTyrLeu 320
Db 34646 GATCCGAAACCCCGAAGACTCTTCGCTTAAAGTACGACGCTCGTTTCAAGACCTTACTTA 34705
Qy 321 LysTyrPheArgPheSerProAspGlyThrLysSerProAspValGluIleGlnLeuAsp 340
Db 34706 AAATACTTCAGATTCACATCCGACGGTGAGAAATCAACAGACGTCGAGATTCGCGTTGAT 34765
Qy 341 GlnProThrMetMetHisAspPheAlaIleThrGluAsnPheValValValProAspGln 360
Db 34766 CAACCAAGATGATCCAGATTTCCGATCTACTGAGAAATTTCTGATGATTTCCAGACCAA 34825
Qy 361 GlnValValPheLysLeuProGluMetIleArgGlyGlySerProValValTyrAspLys 380
Db 34826 CAAGTGGTATTCCGATTCGCGAGATGATAAGAGGTGGTTCTCCCGTGGTTTACGACGAG 34885
Qy 381 AsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSerSerAsnIleLys 400
Db 34886 AAGAAAGAAATCAAGATTCGGGATTTTGAATAAAAAACCGGAAAGATGCTTCGTTCGATTCAA 34945
```

Qy	401	TpIleAspAlaProAspCysPheCysPheHisLeuTrpAsnAlaTrpGluGluProGlu	420
Db	34946	TGGATCGAAGTACCAGATTGTTTGTTCCTCATCTATGAACTCTTGGGAAGAACCGAA	35005
Qy	421	ThrAspGluValValValIleGlySerCysMetThrProProAspSerIlePheAsnGlu	440
Db	35006	ACAGACGAGGTGTCTGTGATTGGATCATGATGACGACCACTGATTCAATTTTCAACGAA	35065
Qy	441	SerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsnLeuLysThrGlyGlu	460
Db	35066	CACGACGAACACTTCAGAGTGTGTTGTCGAGATTAAGCTTAACCTTAAACAGGGGAA	35125
Qy	461	SerThrArgArgProIleIleSerAsnGluAspGlnGlnValAsnLeuGluAlaGlyMet	480
Db	35126	TCAACACGTAGACCGGTTATCTCG-----GAACAAGTTAACTCTCGAAGCTGGTATG	35176
Qy	481	ValAsnArgAsnMetLeuGlyArgIysThrLysPheAlaTyrrLeuAlaLeuAlaGluPro	500
Db	35177	GTCAACCGGAATTTATTATGTTAGAAAAACCGGTATGCTTATCTGGCTTTTAAACGAAACCG	35236
Qy	501	TrpProLysValSerGlyPheAlaLysValAspLeuThrThrGlyGluValLysLysHis	520
Db	35237	TGCGCTTAAAGTGTCCGGTTTCGCTTAAAGTGGACTTATCGACCGAGAGATTTCGAAGATAT	35296
Qy	521	LeuTyrrGlyAsnAsnTyrrGlyGlyGluProLeuPheLeuProGlyGluGlyGlyGlu	540
Db	35297	ATTTCACGAGAGGGGAAATACGGAGGAGAGCCCTCTGTTTCTACCTTCCTCGGTGACGGAGAA	35356
Qy	541	GluAspGluGlyTyrrIleLeuCysPheValHisAspGluLysThrTrpLysSerGluLeu	560
Db	35357	GAAGACCGGAGTTTACATAATGTTGTTTCGTTTCACGACGAGGAGAGTGAATTCGGAATC	35416
Qy	561	GlnIleValAsnAlaValSerLeuGluValGluAlaThrValLysLeuProSerArgVal	580
Db	35417	CAACTCATAAAGCCGCTTAATATGAAGCTTGAAGCTTACCGTAACGCTTCCGTCGAGAGTG	35476
Qy	581	ProTyrrGlyPheHisGlyThrPheIleGlyAlaAspAspLeuAlaLysGlnValVal	599
Db	35477	CCTTACGGTTTCCACGAGACGTTTCATCAGTAGGAAGATTATTCGACGACGAGCTTTG	35533
RESULT 13			
LOCUS	AC013430	95769 bp DNA linear	PLN 26-MAY-2000
DEFINITION	Genomic sequence for Arabidopsis thaliana BAC F3P9 from chromosome I, complete sequence.		
ACCESSION	AC013430	GI:8096768	
VERSION	AC013430.5		
KEYWORDS	HTG.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	1 (bases 1 to 95769)		
AUTHORS	Chao,Q., Brooks,S., Buehler,E., Johnson-Hopson,C., Khan,S., Kim,C., Shinn,P., Altafi,H., Bei,Q., Chin,C., Chiou,J., Choi,E., Conn,L., Conway,A., Gonzales,A., Hansen,N., Howng,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S., Mukharsky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveri,A., Toriumi,M., Vaysberg,M., Yu,G., Federspiel,N.A., Theologis,A. and Ecker,J.R.		
TITLE	I Genomic sequence for Arabidopsis thaliana BAC F3P9 from chromosome I		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 95769)		
AUTHORS	Ecker,J.R.		
TITLE	Direct Submission		
JOURNAL	Submitted (11-NOV-1999) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA		
REFERENCE	3 (bases 1 to 95769)		
AUTHORS	Ecker,J.R.		







QY 48 LeuAsn---ValSerSerAlaLeuHisThrProProAlaLeuHisPheProLysGln--- 65  
 DB 118 TCACACACCATCTGTTCCCTTCAACA-----CTCCACTTCCCCCAACAGTAC 168  
 QY 66 -----SerSerAsnSerProAlaIleValValLysProLysAlaLysGluSer 81  
 DB 169 CAACCAACATCCACATCCACATCCACAGCACACCAACACCCACCCCAATCAAACT 228  
 QY 82 AsnThr----- 83  
 DB 229 ACCACCATCACACACACACCGCCCGGAGAAACCAACCTCTCTCTGACACCAACCA 288  
 QY 84 -----LysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaLeuAspAlaAla 100  
 DB 289 CCATTACCTCAAAAATGGAACTTTCTCCAGAAAGCGCTGCCACGGCCTTGGACCTGGTC 348  
 QY 101 GluGlyPheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSer 120  
 DB 349 GAAACGGCGCTGCTCTCCGACAGCGCAACACCCCGCTCCCAAAAGCGGACCCGAGG 408  
 QY 121 ValGlnIleAlaGlyAsnPheAlaProValAsnGluGlnProValArgArgAsnLeuPro 140  
 DB 409 GTCCAAATCGCCGGGAATTCGCGCGGTGCGGAGCATGCGCGGATCAAGGACTCCCG 468  
 QY 141 ValValGlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyAlaAsnPro 160  
 DB 469 GTGGTCGGAATAATCCCAAAATGCAATTGACGGCGGTGACGTGCGCAACGGTCCCAATCG 528  
 QY 161 LeuHisGluProValThrGlyHisHisPheAspGlyAspGlyMetValHisAlaVal 180  
 DB 529 CTCTAGACCTGTGCGCGGCACCACTTCTTCGACGGCGGCGCATGCTCCACGCGGTG 588  
 QY 181 LysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheVal 200  
 DB 589 AAGTTCACGAACGGCGCCAGCTAGCGCTCGCGCTTCACCGAGACGGCGTCTCTCG 648  
 QY 201 GlnGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyLysHisGlyHis 220  
 DB 649 CAGGAGAATCTCTAGCGCGCGCGGTGTTCCGAAAGGCCATCGGGAGCTCCACGCGCAC 708  
 QY 221 ThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaGlyIleValAspPro 240  
 DB 709 TCCGGCATCGCGCGCTCTCTCTCTACGGCGCGGTCTCTTCGGGCTCGTGTATGGG 768  
 QY 241 AlaHisGlyThrGlyValAlaAsnAlaGlyLeuValTyrPheAsnGlyArgLeuLeuAla 260  
 DB 769 TCCAGGGCATGGCGGTGCGAAGCGCGTCTCGTCTACTTCAACACACCATCTTGGCC 828  
 QY 261 MetSerGluAspAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysThr 280  
 DB 829 ATGTCCGAAGACGATTTACCTTACCACGTGAGAAATCACCCCTTAACGGCGACTTAACACC 888  
 QY 281 ValGlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysVal 300  
 DB 889 GTTGCCCGTGTACGATTCACGGGCGAGCTCAACTCAACAATGATCGCCACCCGAAACTG 948  
 QY 301 AspProGluSerGlyGlyLeuPheAlaLeuSerTyrAspValValLysLysProTyrLeu 320  
 DB 949 GACCCCGTCGACGGCGACCTCCACGGCTCAGCTACGAGTCTATTCAGNAGCTTACCTC 1008  
 QY 321 LysTyrPheArgPheSerProAspGlyThrLysSerProAspValGluIleGlnLeuAsp 340  
 DB 1009 AAGTACTTCCGTTCTTCCCGCCGACGGCGTCAAGTCCCGCGAGCTGGAATCCCGCTGAAG 1068  
 QY 341 GlnProThrMetMetHisPheAlaIleThrGluAsnPheValValValProAspGln 360  
 DB 1069 GAGCCACCATGATGACGATTCGCAATACGGAAATTTCTGCTGCTCCCGACACAG 1128  
 QY 361 GlnValValPheLysLeuProGluMetIleArgGlyGlySerProValValTyrAspLys 380  
 DB 1129 CAGGTGGTCTTCAACTACGAGAGATGATCACCGGGGTGCCCGTGTCTACGACAG 1188  
 QY 381 AsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSerSerAsnIleLys 400

DB 1189 AACAAAACTCTACGGTTTGGGATTCTGCACAAGAAATGCGAAGACCGGAATCGGTGCGG 1248  
 QY 401 TrpIleAspAlaProAspCysPheHisLeuTrpAsnAlaTrpGluGluProGlu 420  
 DB 1249 TGGATCGACGCCCGCGATTGTTCTGCTTCCACCTCTGGAACGCGTGGAGAGGCCGAA 1308  
 QY 421 ThrAspGluValValIleGlySerCysMetThrProProAspSerIlePheAsnGlu 440  
 DB 1309 ACCGAGGAGGTGTGTGTTGGTCTGTCATGACCCCTGCGGACTCCATTTTCAACGAA 1368  
 QY 441 SerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsnLeuLysThrGlyGlu 460  
 DB 1369 TCGCAGGAGAGTTTGAAGAGCGTGTGTGAGATGAAGCTGAACCTTGAGGACCGCAAG 1428  
 QY 461 SerThrArgArgProIleIleSerAsnGluAspGlnGlnValAsnLeuGluAlaGlyMet 480  
 DB 1429 TCCACTCGCGCGCCCATTTATCTCC---GACGCCGAACAGTGAACCTGGAACCGCGCATG 1485  
 QY 481 ValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeuAlaLeuAlaGluPro 500  
 DB 1486 GTGACACAGAAACAAGCTCGGAAGGAAGACCCAGTTTCGGGTATCTGGCTCTGGCGGAGCCC 1545  
 QY 501 TrpProLysValSerGlyPheAlaLysValAspLeuThrThrGlyGluValLysLysHis 520  
 DB 1546 TGGCCCAAGTCTCTGGGCTTTTCGAAAGTTGATTTGCTGAGTGGGGAAGTGAAGAAGTAC 1605  
 QY 521 LeuTyrGlyAspAsnArgTyrGlyGluProLeuPheLeuProGlyGluGlyGlu 540  
 DB 1606 ATGTATGAGAGAGAGAGTTCGGTGGGAGCGCTCTGTTTCTTCCC---AACGCCCAAA 1662  
 QY 541 GluAspGluGlyTyrIleLeuCysPheValHisAspGluLysThrTrpLysSerGluLeu 560  
 DB 1663 GAAGACGATGGTATATTCTGCACTTCGTGCACGACGAGAAAGATGGAATCCGAGCTG 1722  
 QY 561 GlnIleValAsnAlaValSerLeuGluValGluAlaThrValLysLeuProSerArgVal 580  
 DB 1723 CAGATTGTGAATGCCAAAATTTAAAGCTCGAAGCTTCATCAAACTCCCTCTCGTGT 1782  
 QY 581 ProTyrGlyPheHisGlyThrPheIleGlyAlaAspAspLeuAlaGln 597  
 DB 1783 CCCTACGGTTTTCATGGAACTTTTCATTCATCCAAAGGATTTGAGGAACAA 1833  
 RESULT 15  
 LOCUS BD017434 1839 bp DNA linear PAT 27-AUG-2002  
 DEFINITION Transgenic plant using neozanthine cleaving enzyme gene.  
 ACCESSION BD017434  
 VERSION BD017434.1 GI:22558610  
 KEYWORDS JP 2001258579-A/6.  
 SOURCE Vigna unguiculata (cowpea)  
 ORGANISM Vigna unguiculata  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Vigna.  
 REFERENCE 1 (bases 1 to 1839)  
 AUTHORS Iuchi, K., Kobayashi, M. and Shinozaki, K.  
 TITLE Transgenic plant using neozanthine cleaving enzyme gene  
 JOURNAL Patent: JP 2001258579-A 6 25-SEP-2001;  
 THE INSTITUTE OF PHYSICAL AND CHEMICAL RESEARCH  
 COMMENT OS Vigna unguiculata  
 PN JP 2001258579-A/6  
 PD 25-SEP-2001  
 PF 11-JAN-2001 JP 2001003476  
 PI KIYOSHI IUCHI,MASATOMO KOBAYASHI,KAZUO SHINOZAKI PC  
 C12N15/09,A01H5/00,C12N5/10,C12N5/02,C12N15/00,C12N5/00 CC  
 Transgenic plant using neozanthine cleaving enzyme gene PH Key  
 Location/Qualifiers (1)..  
 FT CDS Location/Qualifiers (1)..  
 1..1839  
 /organism="Vigna unguiculata"

```
/mol_type="genomic DNA"
/db_xref="taxon:3917"
BASE COUNT  447 a  595 c  442 g  355 t
ORIGIN

Alignment Scores:
Pred. No.: 178e-148 Length: 1839
Score: 2168.50 Matches: 423
Percent Similarity: 78.44% Conservative: 61
Best Local Similarity: 68.56% Mismatches: 98
Query Match: 68.84% Indels: 35
DB: 6 Gaps: 8

US-09-758-269-6 (1-595) x BD017434 (1-1839)
QY 8 AlaAlaValSerGlyArgTyrLeuGlyGlyAsnHisThrGlnProProLeuSerSerSer 27
DB 7 TCATCAGCTTCAAAACACTTGGTTTAAACGCCACACTCCCATCTCCGCCCTTCAAAAGACCTA 66
QY 28 GlnSerSerAspLeuSerTyrCysSerSerLeuProMetAlaSerArgValThrArgLys 47
DB 67 CCTTCCACATCT--TCTCCCAAAACTTACTTCTTTA-----AGGAAACATCCCTCT 117
QY 48 LeuAen---ValSerSerAlaLeuHisThrProAlaLeuHisPheProLysGln--- 65
DB 118 TCCAAACACCATCAGTGTCTCCCTTCAACA-----CTCCACTTCCCAACACAGTAC 168
QY 66 -----SerSerAsnSerProAlaLeuValValLysProLysAlaLysGluSer 81
DB 169 CAACCAACATCCACATCCACATCCACAGCCACCACCAACACCCCAATCAAAACT 228
QY 82 AsnThr----- 83
DB 229 ACCACCATCACACACACCGCCCGAGGAACCAACCTCTCTCTGACCAACCA 288
QY 84 -----LysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaLeuAspAlaAla 100
DB 289 CCATTACCTCAAAATGGAATCTTCCAGAAAGCGGTGCCACGGCTTGGACCTGGTC 348
QY 101 GluGlyPheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSer 120
DB 349 GAAACGGCGCTGCTCTCGACGAGCGCAACACCGCTCCCAAAACCGCGGACCCGAGG 408
QY 121 ValGlnIleAlaGlyAsnPheAlaProValAsnGluGlnProValArgArgAsnLeuPro 140
DB 409 GTCCAAATTCGGCGGGAATCTCGCGCGGTGCGGAGCATGCCGGATCAAGGACTTCCCG 468
QY 141 ValValGlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyValAsnPro 160
DB 469 GTGGTCGGAAATATCCCAATGCAATTCACGGCGGTGTACGTGCGCAACGGTGCCTCG 528
QY 161 LeuHisGluProValThrGlyHisPhePheAspGlyAspGlyMetValHisAlaVal 180
DB 529 CTCTACGAGCTGTGGCGCGGCACCATCTTTCACGCGCGCATGGTCCACGCGCGTG 588
QY 181 LysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheVal 200
DB 589 AAGTTTCAGAAACGGCGCGCGCATGCTACCGCTCCCGCTTCCAGAGCAGCAGCGTCTCTCG 648
QY 201 GlnGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGlyHis 220
DB 649 CAGGGAATCTCTAGGCGCGCGGTGTTCGGAAGGCCATCGGGGAGCTCCACGGCCAC 708
QY 221 ThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaAlaGlyIleValAspPro 240
DB 709 TCCGCGATCGCGCGCTCTCTCTTCTACGCGCGGTCTCTTCCGGCTCGTTGATGGG 768
QY 241 AlaHisGlyThrGlyValAlaAsnAlaGlyLeuValTyrPheAsnGlyArgLeuLeuAla 260
DB 769 TCCAGGGCATGGGCGTGGCAACCGCGTCTCGTCTACTTCAACCAACCACTCTTGGCC 828
QY 261 MetSerGluAspAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysThr 280
DB 943
```

```
DB 829 ATGTCGGAAGACGATTTTACCCTACCACGTCGAGAATCACCCCTAACGGCGACTTAACACC 888
QY 281 ValGlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysVal 300
DB 889 GTTGGCGCTTACGACTTCAACGGCGAGCTCAACTCAACAATGATGCGCCACCCGAAACTG 948
QY 301 AspProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValLysProTyrLeu 320
DB 949 GACCCCGTCGACGGCGACTCCACGGCTCAGCTACGACGCTCATTCAGAGCCTTACCTC 1008
QY 321 LysTyrPheArgPheSerProAspGlyThrLysSerProAspValGluIleGlnLeuAsp 340
DB 1009 AAGTACTTCCGTTTCTCCCGCAGCGCTCAAGTCCCGCAGCGTGAATCCCTCTGAAG 1068
QY 341 GlnProThrMetMetHisAspPheAlaIleThrGluAsnPheValValProAspGln 360
DB 1069 GAGCCCAACCATGATGACAGATTTCCGCATAACGAGAAATTCGTCGTCGCCGACGAG 1128
QY 361 GlnValValPheLysLeuProGluMetIleArgGlyGlySerProValValTyrAspLys 380
DB 1129 CAGTGTGTCTTCAAACTTACGAGATGATCACCGCGGTCCCCCGTGGTCTACGACAG 1188
QY 381 AsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSerSerAsnIleLys 400
DB 1189 AACAAACCTCACGGTTTGGATTTCTGCACAAGAAATCGAAGGACGCGAATGCGATGCG 1248
QY 401 TrpIleAspAlaProAspCysPheHisLeuTrpAsnAlaTrpGluGluProGlu 420
DB 1249 TGGATCAGCGCGCGGATTTCTGCTTCCACTCTGGAACGCGTGGGAGGAGCCCGAA 1308
QY 421 ThrAspGluValValIleGlySerCysMetThrProProAspSerIlePheAsnGlu 440
DB 1309 ACCGAGAGGTGTGGTGATGGTCTGTCATGCCCTCGCGACTCCATTTTCAACGAA 1368
QY 441 SerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsnLeuLysThrGlyGlu 460
DB 1369 TGCAGGAGAGTTTGAAGAGCGTCTGTACAGATAAGCTGAACCTTGAGACGCGCAAG 1428
QY 461 SerThrArgArgProIleIleSerAsnGluAspGlnGlnValAsnLeuGluAlaGlyMet 480
DB 1429 TCCACTCGCGCGCCCATTTATCTCC--GACGCCGAAACAGTGAACCTTGGAGCCGCGATG 1485
QY 481 ValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeuAlaLeuAlaGluPro 500
DB 1486 GTGAACAGAAACAGCTCGGAAGAGACCCAGTTCGCGTATCTGGCTCTGCGCGAGCC 1545
QY 501 TrpProLysValSerGlyPheAlaLysValAspLeuThrThrGlyGluValLysHis 520
DB 1546 TGGCCCAAGTCTCGGCTTTGCGAAAGTTGATTTGCTGAGTGGGGAAGTGAAGAAGTAC 1605
QY 521 LeuTyrGlyAspAsnArgTyrGlyGluProLeuPheLeuProGlyGlyGlyGlu 540
DB 1606 ATGTATGAGAAGAGAGTTCGGTGGGAGCCTCTGTTTCTTCCC--AACGGCCAAAAA 1662
QY 541 GluAspGluGlyTyrIleLeuCysPheValHisAspGluLysThrTrpLysSerGluLeu 560
DB 1663 GAACAGCATGGGTATATCTGGCATTCGTGCACACGAGAAAGATGGAATTCGAGCTG 1722
QY 561 GlnIleValAsnAlaValSerLeuGluValGluAlaThrValLysLeuProSerArgVal 580
DB 1723 CAGATTGTGAATGCCCAAAATTTAAAGCTCGAAGCTTCCATCAAACTCCCTCTCTGTT 1782
QY 581 ProTyrGlyPheHisGlyThrPheIleGlyAlaAspLeuAlaLysGln 597
DB 1783 CCCTACGTTTTCATGGAACCTTTCATTTCCCAAGGATTTGAGGAACAA 1833
```

Search completed: November 15, 2003, 15:02:42

Job time : 6417 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 15, 2003, 09:19:43 ; Search time 447 Seconds  
(without alignments)  
3617.368 Million cell updates/sec

Title: US-09-758-269-6  
Perfect score: 3150  
Sequence: 1 MASFTATAAASGRWLGNGHT.....VPYGFHGTGICADDLAKQVV 599

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5	
Fgapop 6.0	Fgapext 7.0	
Delop 6.0	Delext 7.0	

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cpn2.1/USPTO spool/US09758269/runat 14112003 135755 28735/app query.fasta\_1.775  
-DB=N Geneseq 15Jun03 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09758269 @CEN 1.1 312 @runat 14112003 135755 28735 -NCPU=6 -ICPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 19Jun03:\*

- 1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*
- 2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*
- 3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*
- 4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*
- 5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*
- 6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*
- 7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*
- 8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*
- 9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*
- 10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*
- 11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*
- 12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*
- 13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*
- 14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*
- 15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*
- 16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*
- 17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*
- 18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*
- 19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*
- 20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*
- 21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*
- 22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*
- 23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*
- 24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*
- 25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	3150	100.0	1800	22	AAD09396	Arabidopsis thalia
2	2280.5	72.4	1818	22	AAD09401	Lycopersicon escul
3	2168.5	68.8	1839	22	AAD09399	Vigna unguiculata
4	1991	63.2	1752	22	AAD09394	Arabidopsis thalia
5	1930	61.3	1815	22	AAD09400	Zea mays neoxanthi
6	1663.5	52.8	1734	22	AAD09398	Arabidopsis thalia
7	969	30.8	1788	21	AAC42989	Arabidopsis thalia
8	969	30.8	1788	22	AAD09395	Arabidopsis thalia
9	969	30.8	1788	24	ABZ13639	Arabidopsis thalia
10	960.5	30.5	1950	22	AAF77206	cDNA encoding sunf
11	939	29.8	1777	21	AAC36083	Arabidopsis thalia
12	938	29.8	1617	22	AAD09397	Arabidopsis thalia
13	937	29.7	1617	22	AAD09402	Arabidopsis thalia
14	608	19.3	443	21	AAC56678	Eucalyptus grandis
15	592	18.8	492	22	AAC82706	Rice abscisic acid
16	556	17.7	491	21	AAC57157	Pinus radiata tran
17	522	16.6	393	25	ABX20484	Human GDP-mannose
18	511	16.2	372	21	AAC56695	Eucalyptus grandis
19	494.5	15.7	4403765	22	AA199683	Mycobacterium tube
20	494.5	15.7	4411529	22	AA199682	Mycobacterium tube
21	475.5	15.1	4403765	22	AA199683	Mycobacterium tube
22	475.5	15.1	4411529	22	AA199682	Mycobacterium tube
23	463	14.7	325	21	AAC56548	Eucalyptus grandis
24	415	13.2	386	21	AAC57162	Pinus radiata tran
25	409.5	13.0	340	21	AAC56520	Eucalyptus grandis
26	401	12.7	329	21	AAC57165	Pinus radiata tran
27	398	12.6	326	21	AAC57167	Pinus radiata tran
28	371.5	11.8	398	21	AAC57145	Pinus radiata tran
29	369	11.7	320	25	ABX20813	Human GDP-mannose
30	337.5	10.7	238	21	AAC57169	Pinus radiata tran
31	330.5	10.5	567	21	AAC56546	Eucalyptus grandis
32	322.5	10.2	1713	22	AAD19635	Arabidopsis thalia
33	322	10.2	470	24	ABL57537	Apoptosis inhibito
34	311	9.9	246	21	AAC57168	Pinus radiata tran
35	311	9.9	412	21	AAC57164	Pinus radiata tran
36	307.5	9.8	1791	22	AAD19634	Arabidopsis thalia
37	290	9.2	308	21	AAC57142	Pinus radiata tran
38	286.5	9.1	395	21	AAC57117	Pinus radiata tran
39	279	8.9	534	22	AAH44254	Physcomitrella pat
40	272.5	8.7	190	21	AAC57163	Pinus radiata tran
41	272.5	8.7	501	21	AAC43547	Zea mays DNA fragm
42	266.5	8.5	508	21	AAC56511	Eucalyptus grandis
43	260	8.3	3111	21	AAH53888	Beta, Beta-caroten
44	256	8.1	1773	22	AAH42170	Nucleotide sequenc
45	252.5	8.0	447	21	AAC56480	Eucalyptus grandis

ALIGNMENTS

RESULT 1  
AAD09396  
ID AAD09396 standard; cDNA; 1800 BP.  
XX AAD09396;  
AC AAD09396;  
XX  
DT 10-SEP-2001 (first entry)  
XX  
DE Arabidopsis thaliana neoxanthin cleavage enzyme, AtNCED3 cDNA.  
XX  
KW Neoxanthin cleavage enzyme; AtNCED3; abscisic acid; ABA; herbicide;  
KW stress tolerance; transgenic plant; plant breeding; antisense-therapy;  
KW plant growth protectant; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
FH Key Location/Qualifiers

```
FT CDS 1..1800
FT /+tag= a
FT /product= "Arabidopsis thaliana AtNCED3 protein"
XX
XX EP1116794-A2.
XX
XX 18-JUL-2001.
XX
XX 11-JAN-2001; 2001EP-0300218.
XX
XX 13-JAN-2000; 2000JP-0010056.
XX
XX 11-JAN-2001; 2001JP-0003476.
XX
XX (RIKE ) RIKEN KK.
XX
XX Iuchi S, Kobayashi M, Shinozaki K;
XX WPI; 2001-400081/43.
XX P-PSDB; AAE04784.
XX
XX A DNA encoding a protein with a neoxanthin cleavage activity for
XX producing transgenic plants with improved or decreased stress tolerance
XX
XX
XX Claim 3; Page 32-36; 101pp; English.
XX
XX The invention relates to neoxanthin cleavage enzymes and their
XX corresponding cDNA molecules. Neoxanthin cleavage enzyme plays a key
XX role in endogenous abscisic acid (ABA) biosynthesis under drought stress.
XX Neoxanthin cleavage enzyme is used for improving stress tolerance in a
XX plant when expressed in a plant cell. The invention also relates to
XX methods for increasing or decreasing stress tolerance in a plant by
XX introducing the DNA into the plant, and a transgenic plant into which a
XX neoxanthin cleavage enzyme is introduced. The improvement of stress
XX tolerance in plants is useful, for example in plant breeding. Neoxanthin
XX cleavage enzyme genes are useful for producing transgenic plants. An arid
XX land can be improved by growing transformat weed for several years and
XX then removing the weed by specifically lowering stress tolerance in the
XX weed by inducing an inducible promoter. The present cDNA sequence encodes
XX Arabidopsis thaliana neoxanthin cleavage enzyme, AtNCED3 protein.
XX The AtNCED3 cDNA is obtained from an Arabidopsis plant-derived cDNA
XX library using a cDNA of the CPRD65 (Cowpea Responsive to Dehydration)
XX gene isolated from cowpea plant as a probe.
XX
XX SQ Sequence 1800 BP; 458 A; 464 C; 439 G; 439 T; 0 other;

Alignment Scores:
Pred. No.: 3.92e-298 Length: 1800
Score: 3150.00 Matches: 599
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0

US-09-758-269-6 (1-599) x AAD09396 (1-1800)

Qy 1 MetAlaSerPheThrAlaThraAlaValSerGlyArgTrpLeuGlyGlyAsnHisThr 20
Db 1 ATGGCTTCTTTCACGGCAACGGCTGCGGTTCTGGAGATGGCTTGGTGGCAATCACT 60
Qy 21 GlnProLeuSerSerGlnSerSerAspLeuSerTyrCysSerSerLeuProMet 40
Db 61 CAGCGCCCATATGCTGCTCTCTCAAGCTCCGACCTTGAGTTAGTCTCTTACCTATG 120
Qy 41 AlaSerArgValThrArgLysLeuAsnValSerSerAlaLeuHisThrProAlaLeu 60
Db 121 GCCAGTCGTGTACACAGTAAGCTCAATGTTTCATCTGGCGCTTCACACTCTCCAGCTTT 180
Qy 61 HisPheProLysGlnSerSerAsnSerProAlaIleValValLysProLysAlaLysGlu 80
Db 181 CATTTCCCTAAGCAATCATCAAACTCTCCCGCCATTGTTGTTAAGCCCAAGCCAAAGAA 240
Qy 81 SerAenThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaLeuAspAlaAla 100
```

```
Db 241 TCACACACTAAACAGATGAATTTGTTCCAGAGAGCGCGCGGAGGTTGGAGCGCGCG 300
Qy 101 GluGlyPheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSer 120
Db 301 GAGGGTTTCTTGTGACGACAGAGAGCTACCCGCTTCTTAAACCGCTAGTCCTAGT 360
Qy 121 ValGlnIleAlaGlyAsnPheAlaProValAsnGluGlnProValArgArgAsnLeuPro 140
Db 361 GTTCAGATCCGCCGAAATTTTGCTCCGGTGAATGAACAGCCCGCTCCGGCGTAAATCTTC 420
Qy 141 ValValGlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyAlaAsnPro 160
Db 421 GTGGTCGAAACCTTCCGATTCATCAAGAGAGTGTATGTGCGCAACGAGCTAACCCA 480
Qy 161 LeuHisGluProValThrGlyHisHisPheAspGlyAspGlyMetValHisAlaVal 180
Db 481 CTTACGAGCGGTGACAGGTCACTTCTTCGAGGAGACGCTATGTTTACGCGGTC 540
Qy 181 LysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheVal 200
Db 541 AAATTCGAACACCGGTTCCAGTAGCTAGCTTGCCTGCGGTTTACTCAGACTAACCGGTTTGT 600
Qy 201 GlnGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGlyHis 220
Db 601 CAGGAAGCTCAATTTGGTGCACCGGTTTCCCAAGCCATCGGTGAGCTTCACGGCCAC 660
Qy 221 ThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaAlaGlyIleValAspPro 240
Db 661 ACCGTTATTCCCGACTCATGCTATTCTACGCCAGAGCTGCAGCGGTATAGTCGACCCG 720
Qy 241 AlaHisGlyThrGlyValAlaAsnAlaGlyLeuValTyrPheAsnGlyArgLeuLeuAla 260
Db 721 GCACACGAAACCGGTAGCTAAACCGGTTTGGTCTATTTCATTAATGCGCGGTTATTGGCT 780
Qy 261 MetSerGluAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysThr 280
Db 781 ATGTCGGAGAGATGATTACCTTACCAAGTTCCAGATCACTCCCAATGGAGATTAAAAACC 840
Qy 281 ValGlyArgPheAspPheAspGlyGlnLeuGlySerThrMetIleAlaHisProLysVal 300
Db 841 GTTGGTCCGTTTCGATTTTGTATGGACAATTAGAATCCACAATGATTGCCCCCAGAAAGTC 900
Qy 301 AspProLysSerGlyGluLeuPheAlaLeuSerTyrAspValValSerLysProTyrLeu 320
Db 901 GACCCGGAATCCGGTGAACCTTTCGCTTAAAGCTTACGACGCTGCTTCAAAGCCTTACCTA 960
Qy 321 LysTyrPheArgPheSerProAspGlyThrLysSerProAspValGluIleGlnLeuAsp 340
Db 961 AATACTTCCGATTTCTCACCAGCGGAACCTAATCACCAGGAGCTCGAGATTACGTTGAT 1020
Qy 341 GlnProThrMetMetHisAspPheAlaIleThrGluAsnPheValValProAspGln 360
Db 1021 CAGCCAAACGATGATGCACGATTTCCGATTTACAGAGAATTCGTCGTCGTACCTGACCCAG 1080
Qy 361 GlnValValPheLysLeuProGluMetIleArgGlyGlySerProValValTyrAspLys 380
Db 1081 CAAAGTCGTTTCAAGCTCCCGGAGATGATCCGCGGTGGGTCTCCGGTGGTTTACGACAAG 1140
Qy 381 AsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSerSerAsnIleLys 400
Db 1141 AACAAAGTCGCAAGATTCGGGATTTTAGACAAATACCCGAGATTTCATCGAACATTAAAG 1200
Qy 401 TrpIleAspAlaProAspCysPheCysPheHisLeuTrpAsnAlaTrpGluGluProGlu 420
Db 1201 TGGATTGATGTCCTCCAGATTGCTTCTGCTTCCATCTCTGGAACGCTTGGGAAGAGCCAGAA 1260
Qy 421 ThrAspGluValValIleGlySerCysMetThrProProAspSerIlePheAsnGlu 440
Db 1261 ACAGATGAAGTCGCTGATAGGCTCTGTATGACTCCACAGACTCAATTTTCAACGAG 1320
Qy 441 SerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsnLeuLysThrGlyGlu 460
```

```
Db 1321 TCTGACGAGAAATCTCAAGAGTGCTCTGTCTGAAATCCGCTGAATCTCAAAACCGGTGAA 1380
Qy 461 SerThrArgArgProIleIleSerAsnGluAspGlnGlnValAsnLeuGluAlaGlyMet 480
Db 1381 TCAACTCGCGCTCGGATCATCTCCAACGAAGATCAACAAGTCAACCTCGAAGCAGGATG 1440
Qy 481 ValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrlleuAlaLeuAlaGluPro 500
Db 1441 GTCAACAGAAACATGCTCGGCGGTAAACCAAAATTCGCTTACTTGCGCTTTAGCCGAGCG 1500
Qy 501 TrpProLysValSerGlyPheAlaLysValAspLeuThrThrGlyGluValLysLysHis 520
Db 1501 TGGCCTAAAGTCTCAGGATTCGTAAGTTGATCTCAGTCTGAGAGAACTTGAAGAAAT 1560
Qy 521 LeuTyrlGlyAspAsnArgTyrlGlyGluProLeuPheLeuProGlyGlyGlyGlu 540
Db 1561 CTTTACGGCGATAACCGTTACGAGGAGAGCGCTCTGTTTCTCCCGAGAGAGAGAG 1620
Qy 541 GluAspGluGlyTyrlleuLeuCysPheValHisAspGluLysThrTrpLysSerGluLeu 560
Db 1621 GAAGACGAAGAGATACATCTCTGTTTCGTTCAAGCAGAGAGACATCGAAATTCGGAGTTA 1680
Qy 561 GlnIleValAsnAlaValSerLeuGluValGluAlaThrValLysLeuProSerArgVal 580
Db 1681 CAGATAGTTAACCGCTTAGCTTAGGTTGAGGTTGAGCAACCGTTAAACTCCGTCAGGGTT 1740
Qy 581 ProTyrlGlyPheHisGlyThrPheIleGlyAlaAspAspLeuAlaLysGlnValVal 599
Db 1741 CCGTACGGATTTTCACGGTACATTCATCGGAGCGGATGATTTGCGAAGCAGGTCGTG 1797

RESULT 2
AAD09401
ID AAD09401 standard; cDNA; 1818 BP.
XX
AC AAD09401;
XX
DT 10-SEP-2001 (first entry)
XX
DE Lycopersicon esculentum neoxanthin cleavage enzyme, LeNCEd1 cDNA.
XX
KW Tomato; neoxanthin cleavage enzyme; LeNCEd1; abscisic acid; ABA;
KW stress tolerance; transgenic plant; plant breeding; antisense-therapy;
KW plant growth protectant; herbicide; ss.
XX
OS Lycopersicon esculentum.
XX
FH Key Location/Qualifiers
FT CDS 1..1818
FT /*tag= a
FT /product= "Lycopersicon esculentum LeNCEd1 protein"
XX
FN EP1116794-A2.
XX
PD 18-JUL-2001.
XX
PF 11-JAN-2001; 2001EP-0300218.
XX
PR 13-JAN-2000; 2000JP-0010056.
PR 11-JAN-2001; 2001JP-0003476.
XX
XX (RIKE ) RIKEN KK.
XX
XX Iuchi S, Kobayashi M, Shinozaki K;
XX
XX WPI: 2001-400081/43.
XX P-PSDB; AAE04789.
XX
XX A DNA encoding a protein with a neoxanthin cleavage activity for
XX producing transgenic plants with improved or decreased stress tolerance
XX
XX Claim 3; Page 67-71; 101pp; English.
XX
```

```
CC The invention relates to neoxanthin cleavage enzymes and their
CC corresponding cDNA molecules. Neoxanthin cleavage enzyme plays a key
CC role in endogenous abscisic acid (ABA) biosynthesis under drought stress.
CC Neoxanthin cleavage enzyme is used for improving stress tolerance in a
CC plant when expressed in a plant cell. The invention also relates to
CC methods for increasing or decreasing stress tolerance in a plant by
CC introducing the DNA into the plant, and a transgenic plant into which a
CC neoxanthin cleavage enzyme is introduced. The improvement of stress
CC tolerance in plants is useful, for example in plant breeding. Neoxanthin
CC cleavage enzyme genes are useful for producing transgenic plants. An arid
CC land can be improved by growing transformant weed for several years and
CC then removing the weed by specifically lowering stress tolerance in the
CC weed by inducing an inducible promoter. The present cDNA sequence encodes
CC Lycopersicon esculentum neoxanthin cleavage enzyme, LeNCEd1 protein
CC related to the invention.
XX
SQ Sequence 1818 BP; 569 A; 383 C; 369 G; 497 T; 0 other;

Alignment Scores:
Pred. No.: 5,17e-213 Length: 1818
Score: 2280.50 Matches: 435
Percent Similarity: 82.04% Conservative: 63
Best Local Similarity: 71.66% Mismatches: 90
Query Match: 72.40% Indels: 19
DB: 22 Gaps: 6

US-09-758-269-6 (1-599) x AAD09401 (1-1818)
Qy 5 ThrAlaThrAlaAlaValSerGlyArgTrpLeuGlyGlyAsnHisThrGlnProLeu 24
Db 7 ACTACTACTTCCATGCGCCAAATACATACATGATTTTTCATCACTCACTTCTCTACTCAAAAT 51
Qy 25 SerSerSerGlnSerSerAspLeuSerTyrlCysSer-----SerLeuProMetAla 41
Db 52 TCATGTCATCATCAAGAGAGTTGGTTTTCATCACTCACTTCTCTACTCAAAAT 111
Qy 42 SerArgValThrArgLysLeuAsnValSerSerAlaLeuHisThrProAlaLeuHis 61
Db 112 CAACATATAGGCAAGTCTCAACATTAATTCCTCTCTTCAAGTCCACCTATATCTAT 171
Qy 62 PheProLysGlnSerSerAsn-----SerProAla-----IleValLysProLys 77
Db 172 TTTCTTAACCAATCTTCAAAATTTATCAACACCAAGAAATATACAAATTTTCACACCCAAA 231
Qy 78 AlaLysGluSerAsn-----ThrLysGlnMetAsnLeuPheGlnArgAla 92
Db 232 CAAGAAACAACTCTCTCTCTTCTCACTTCCAAAGTGGAAATTTAGTCGAGAAAGA 291
Qy 93 AlaAlaAlaLeuAspAlaAlaGluGlyPheLeuValSerHisGluLysLeuHisPro 112
Db 292 GCAGCAATGGCTTTAGATGCTGTAGAAAGTCTTTAACTAAACATGAACCTTGACACCT 351
Qy 113 LeuProLysThrAlaAspProSerValGlnIleAlaGlyAsnPheAlaProValAsnGlu 132
Db 352 TTGCCGAAACAGCGCACCCACGAGTCCAGATTTCTGGGAAATTTTGCTCCGCTACCGAA 411
Qy 133 GlnProValArgAsnLeuProValValGlyLysLeuProAspSerIleLysGlyVal 152
Db 412 AATCAGTCTGTCATCTCTCCGTCACCGGAAATATACCAATGTTGTTCAAGGCGTT 471
Qy 153 TyrValArgAsnGlyAlaAsnProLeuHisGluProValThrGlyHisPhePheAsp 172
Db 472 TACGTTGAAACGGAGGCTAACCTCTTTTGAACCAACCGCGGACACCATTTCTTCGAC 531
Qy 173 GlyAspGlyMetValHisAlaValLysPheGluHisGlySerAlaSerTyrlAlaCysArg 192
Db 532 GGCAGCGTATGTTGATGCTTCAAGTTCATTAATAAATGGGTGCGCTAGTTAGCTTGGCGT 591
Qy 193 PheThrGlnThrAsnArgPheValGlnGluArgGlnLeuGlyArgProValPheProLys 212
Db 592 TTCACCTGAACAGAGAGGCTTTGTTCAAGAAAGAGCTTTGGGTGCGCTTGTTCCTCTAAA 651
Qy 213 AlaIleGlyGluLeuHisGlyHisThrGlyIleAlaArgLeuMetLeuPheTyrlAlaArg 232
```

```
CC The invention relates to neoxanthin cleavage enzymes and their
CC corresponding cDNA molecules. Neoxanthin cleavage enzyme plays a key
CC role in endogenous abscisic acid (ABA) biosynthesis under drought stress.
CC Neoxanthin cleavage enzyme is used for improving stress tolerance in a
CC plant when expressed in a plant cell. The invention also relates to
CC methods for increasing or decreasing stress tolerance in a plant by
CC introducing the DNA into the plant, and a transgenic plant into which a
CC neoxanthin cleavage enzyme is introduced. The improvement of stress
CC tolerance in plants is useful, for example in plant breeding. Neoxanthin
CC cleavage enzyme genes are useful for producing transgenic plants. An arid
CC land can be improved by growing transformant weed for several years and
CC then removing the weed by specifically lowering stress tolerance in the
CC weed by inducing an inducible promoter. The present cDNA sequence encodes
CC Lycopersicon esculentum neoxanthin cleavage enzyme, LeNCEd1 protein
CC related to the invention.
XX
SQ Sequence 1818 BP; 569 A; 383 C; 369 G; 497 T; 0 other;
```



DB: 22 Gaps: 8  
US-09-758-269-6 (1-599) x AAD09399 (1-1839)  
Qy 8 AlaAlaValSerGlyArgTyrLeuGlyGlyAsnHisThrGlnProProLeuSerSerSer 27  
Db 7 TCATCAGCTTCAACACACTTGGTTTAACGCCACACTCCCATCTCCCCCTTCAAGACCTA 66  
Qy 28 GlnSerSerAspLeuSerTyrCysSerLeuProMetAlaSerArgValThrArgLys 47  
Db 67 CCTTCCACATCT--TCTCCACAACTTACTTCTCTTA-----AGGAAACATCTCT 117  
Qy 48 LeuAsn---ValSerSerAlaLeuHisThrProAlaLeuHisPheProLysGln--- 65  
Db 118 TCCAACACCATCACAATGTTCCCTTCAACA-----CTCCACTTCCCAACAGTAC 168  
Qy 66 -----SerSerAsnSerProAlaIleValValLysProLysAlaLysGluSer 81  
Db 169 CAACCAACATCCACATCCACATCCACAGCCACCAACACACCCCAATCAAACT 228  
Qy 82 AsnThr----- 83  
Db 229 ACCACCATCACCACCAACACCGCCAGCGAAACCAACCTCTCTCTGACACCAACCA 288  
Qy 84 -----LysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaLeuAspAla 100  
Db 289 CCATTACCTCAAAATGGAACTTTCTCCAGAAAGCGCTGCCACGCGCTTGGACCTGTC 348  
Qy 101 GluGlyPheLeuValSerHisGlyLysLeuHisProLeuProLysThrAlaAspProSer 120  
Db 349 GAAACGGCGCTCTCTCGCACGAGCGCAACACCGCTCCCAACAGCGCGACCCGAGG 408  
Qy 121 ValGlnIleAlaGlyAsnPheAlaProValAsnGluGlnProValArgArgAsnLeuPro 140  
Db 409 GTCCAAATCGCCGGAACTTCGCGCGGTCCGCGAGCATGCCCGCATCAAGACTCCCG 468  
Qy 141 ValValGlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyAlaAsnPro 160  
Db 469 GTGGTCGGAATAATCCCAATATGCAATGACGGCGGTGACGCGCAACGGTGCCAAATCG 528  
Qy 161 LeuHisGluProValThrGlyHisPhePheAspGlyAspGlyMetValHisAlaVal 180  
Db 529 CTCTACGAGCTTGTGGCGGCGCACCATCTCTTCAGCGCGCGCATGTCACCGCGGTG 588  
Qy 181 LysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheVal 200  
Db 589 AAGTTCCAGAACGGCGCCGACGCTACGCTCCGCTTCCACGACGACGAGCTCTCTCG 648  
Qy 201 GlnGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGlyHis 220  
Db 649 CAGGAGAAATCTCTAGCGCGCGCGGTGTTCCCGAAGGCGCATCGGGGAGCTCCACGCGCAC 708  
Qy 221 ThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaAlaGlyIleValAspPro 240  
Db 709 TCCGCGATCGCGCGCTCTCTCTTCACGCGCGCGGTCTCTTCGCGGTCTGTTGATGGG 768  
Qy 241 AlaHisGlyThrGlyValAlaAlaAlaGlyLeuValTyrPheAsnGlyArgLeuLeuAla 260  
Db 769 TCCAGGCGATGGGCGGTGGGACCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 828  
Qy 261 MetSerGluAspAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysThr 280  
Db 829 ATGTCCGAAGACGATTTACCTTACCACGTCGAGATCAACCTTAAACGCGACTTAAACACC 888  
Qy 281 ValGlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysVal 300  
Db 889 GTTGCGCGTTAGCATTTCAACGGGAGCTCAACTCAACAATGATCGCCACCCCGAACTG 948  
Qy 301 AspProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValSerLysProTyrLeu 320  
Db 949 GACCCCGTCGCGCGGACCTCCACGCGCTCAGCTACGAGCTCATTCAGAACGCTTACCTC 1008  
Qy 321 LysTyrPheArgPheSerProAspGlyThrLysSerProAspValGluIleGlnLeuAsp 340

Db 1009 AAGTACTTCCGTTTCTCCCGACGCGCTCAAGTCCCGACGAGTGGAAATCCCCCTGAAG 1068  
Qy 341 GlnProThrMetMetHisAspPheAlaIleThrGluAsnPheValValProAspGln 360  
Db 1069 GAGCCACCATGATGACGATTTCCCATACGGAATTTCTGCTGCTCCCGACCAAG 1128  
Qy 361 GlnValValPheLysLeuProGluMetIleArgGlyGlySerProValValTyrAspLys 380  
Db 1129 CAGGTGGTCTTCAAACTAACCGAGATGATACCGCGCGGTCCCGCTGTCTACGACAAG 1188  
Qy 381 AsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSerSerAsnLys 400  
Db 1189 AACAAACCTCACGGTTTGGGATTTCTGCACAGAATGCGAAGACCGAAATCGATGCGG 1248  
Qy 401 TrpIleAspAlaProAspCysPheHisLeuTrpAsnAlaTrpGluLysProGlu 420  
Db 1249 TGGATCGACGCGCGGATTTCTGCTTCCACCTCTGGAACGCTGGAGGAGGAGCCGAA 1308  
Qy 421 ThrAspGluValValIleGlySerCysMetThrProProAspSerIlePheAsnGlu 440  
Db 1309 ACCGAGGAGGTGGTGTGTTGGTCTCTGATGACCCCTGCGGACTCCATTTTCAACGAA 1368  
Qy 441 SerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsnLeuLysThrGlyGlu 460  
Db 1369 TCGGAGGAGAGTTGAAGAGCGTGTGTGTCAGAGATAAGCTGAACCTTGAGACCGGCAAG 1428  
Qy 461 SerThrArgArgProIleIleSerAsnGluAspGlnGlnValAsnLeuGluAlaGlyMet 480  
Db 1429 TCCACTCGCGCGCCCATTTATCTCC---GACCGCCGAACGAAGTGAACCTGGAAGCGCGCATG 1485  
Qy 481 ValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeuAlaLeuAlaGluPro 500  
Db 1486 GTGAACAGAAACAAGCTCGGAAGGAAGACCCAGTTCCGCTATCTGCTCTGCGGAGCCC 1545  
Qy 501 TrpProLysValSerGlyPheAlaLysValAspLeuThrThrGlyGluValLysLysHis 520  
Db 1546 TGGCCCAAGTCTCGGGCTTTGCGAAAGTTGATTTGCTGAGTGGGGAAGTGAAGAAGTAC 1605  
Qy 521 LeuTyrGlyAspAsnArgTyrGlyGluProLeuPheLeuProGlyGlyGlyGlu 540  
Db 1606 ATGTATGAGAAGAGAGTTCGTTGGGAGGCTCTGTTTCTTCCC---AACGCGCAAAAC 1662  
Qy 541 GluAspGluGlyTyrIleLeuCysPheValHisAspGluLysThrTrpLysSerGluLeu 560  
Db 1663 GAAGCAGTGGTATATTCTGCTGCTGTCACGAGAGAAAGAAATGGAATCCGAGCTG 1722  
Qy 561 GlnIleValAsnAlaValSerLeuGluValGluAlaThrValLysLeuProSerArgVal 580  
Db 1723 CAGATTGTGAATGCCAAAATTTAAAGCTCGAAGCTTCCATCAAACCTCCCTCTCGTGT 1782  
Qy 581 ProTyrGlyPheHisGlyThrPheIleGlyAlaAspLeuAlaLysGln 597  
Db 1783 CCTACGGTTCATGGAACCTTTCATTCATTCCTCAAGGATTTGAGGAAACNA 1833  
RESULT 4  
AAD09394  
ID AAD09394 standard; cDNA; 1752 BP.  
XX AAD09394;  
AC AAD09394;  
XX 10-SEP-2001 (first entry)  
XX Arabidopsis thaliana neoxanthin cleavage enzyme, AtNCED1 cDNA.  
XX Neoxanthin cleavage enzyme; AtNCED1; abscisic acid; ABA; herbicide;  
KW stress tolerance; transgenic plant; plant breeding; antisense-therapy;  
KW plant growth protectant; ss.  
XX Arabidopsis thaliana.  
OS Arabidopsis thaliana.  
FH Key Location/Qualifiers  
FT CDS 1..1752



QY 474 ValAsnLeuGluAlaGlyMetValAsnArgAsnMetLeuGlyArgLysThrLysPheAla 493  
 DB 1363 GTAAATTTAGATTGGTATGGTTAAACCGAATAGGAAAGAAACCCGGTTTCGGC 1422  
 QY 494 TyrLeuAlaLeuAlaGluProTyrProLysValSerGlyPheAlaLysValAspLeuThr 513  
 DB 1423 TTTTGGCTATGCTTATCTTGGCCAAAGTTTCGGTTTCGTAAGTTCGATCTTTC 1482  
 QY 514 ThrGlyGluValLysLysHisLeuTyrGlyAspAsnArgTyrGlyGluProLeuPhe 533  
 DB 1483 ACCGGTGAGATGAATAATATATTTACGGCGGTGAGAAATATCGCGCGCAACCGTTTTC 1542  
 QY 534 LeuProGlyGluGlyGly-----GluGluAspGluGlyTyrIleLeuCysPhe 549  
 DB 1543 TTGCGCGCAACTCCCGTAACGCGGAAGAAATGAAGATGACGGTTTATATTTGTAC 1602  
 QY 550 ValHisAspGluLysThrTrpLysSerGluLeuGlnIleValAsnAlaValSerLeuGlu 569  
 DB 1603 GTTCATGACGAAGAAACAAAGACATCAGAGCTTCAGATTATTAAACGCTTAAATTTAAAG 1662  
 QY 570 ValGluAlaThrValLysLeuProSerArgValProTyrGlyPheHisGlyThrPheIle 589  
 DB 1663 CTTGAAGCTACGATTAAACTACCGTCTAGTACCGTATGGGTTTCATGCGCATTTG 1722  
 QY 590 GlyAlaAspAspLeuAlaLysGlnVal 598  
 DB 1723 GATTCCGATGAATCTCGTTGATCAATTA 1749

## RESULT 5

AAD09400

ID AAD09400 standard; cDNA; 1815 BP.

XX AAD09400;

DT 10-SEP-2001 (first entry)

XX Zea mays neoxanthin cleavage enzyme, VP14 cDNA.

DE Maize; neoxanthin cleavage enzyme, VP14 cDNA.

KW Maize; neoxanthin cleavage enzyme; VP14; abscisic acid; ABA; herbicide;  
 KW stress tolerance; transgenic plant; plant breeding; antisense-therapy;  
 KW plant growth protectant; ss.

XX Zea mays.

XX Key Location/Qualifiers

FH 1..1815

FT /tag= a

FT /product= "Zea mays VP14 protein"

PN EP1116794-A2.

XX 18-JUL-2001.

XX 11-JAN-2001; 2001EP-0300218.

XX 13-JAN-2000; 2000JP-0010056.

XX 11-JAN-2001; 2001JP-0003476.

XX (RIKE ) RIKEN KK.

XX Iuchi S, Kobayashi M, Shinozaki K;

XX WPI; 2001-400081/43.

XX P-PSDB; AAE04788.

PT A DNA encoding a protein with a neoxanthin cleavage activity for  
 PT producing transgenic plants with improved or decreased stress tolerance  
 PT .  
 XX Claim 3; Page 60-64; 101pp; English.  
 XX The invention relates to neoxanthin cleavage enzymes and their

CC corresponding cDNA molecules. Neoxanthin cleavage enzyme plays a key  
 CC role in endogenous abscisic acid (ABA) biosynthesis under drought stress.  
 CC Neoxanthin cleavage enzyme is used for improving stress tolerance in a  
 CC plant when expressed in a plant cell. The invention also relates to  
 CC methods for increasing or decreasing stress tolerance in a plant by  
 CC introducing the DNA into the plant, and a transgenic plant into which a  
 CC neoxanthin cleavage enzyme is introduced. The improvement of stress  
 CC tolerance in plants is useful, for example in plant breeding. Neoxanthin  
 CC cleavage enzyme genes are useful for producing transgenic plants. An arid  
 CC land can be improved by growing transformant weed for several years and  
 CC then removing the weed by specifically lowering stress tolerance in the  
 CC weed by inducing an inducible promoter. The present cDNA sequence encodes  
 CC Zea mays neoxanthin cleavage enzyme, VP14 protein related to the  
 CC invention.

XX  
 SQ Sequence 1815 BP; 266 A; 689 C; 604 G; 256 T; 0 other;

## Alignment Scores:

Pred. No.: 1.05e-178 Length: 1815  
 Score: 1930.00 Matches: 375  
 Percent Similarity: 75.68% Conservative: 73  
 Best Local Similarity: 63.34% Mismatches: 122  
 Query Match: 61.27% Indels: 22  
 DB: 22 Gaps: 8

US-09-758-269-6 (1-599) x AAD09400 (1-1815)

QY 22 ProProLeuSerSerSerGlnSerSerAspLeuSerTyrCysSerSerLeuProMetAla 41  
 DB 49 CCGGCCCGGTCAGGGCCCGGGCTCC-----AATTCGTAGGTTCTCG 93  
 QY 42 SerArgValThrArgLysLeuAsnValSerSerAlaLeuHisThrPro-----Pro 58  
 DB 94 CCGCGCGCGTCAGCTCGCGCCCGCGAGTCCGTCAGGCGCGTTCACAAAGCCC 153  
 QY 59 AlaLeuHisPheProLysGlnSerSerAsnSerProAlaIleValVal-----74  
 DB 154 GTCGCGGACCTGCCTCGCGCGCTCCAGGAAGCCCGCCCATTTGCCGTCAGGCGCAGCC 213  
 QY 75 ---LysProLysAlaLysGluSerAsnThrLysGlnMetAsnLeuPheGlnArgAla 93  
 DB 214 CGCGCCCGAGAAAGCGAGGGCGGCAAGAGAGAGTCAACTGTTTCAGGCGCGCGCG 273  
 QY 94 AlaAlaAlaLeuAspAla---AlaGluGlyPheLeuValSer---HisGluLysLeuHis 111  
 DB 274 CGCGCCCGCGTCGACGCGTTCGAGGAAGGTTGTTGGCCAAAGTCTCTGAGGCGCCCCAC 333  
 QY 112 ProLeuProLysThrAlaAspProSerValGlnIleAlaGlyAsnPheAlaProValAsn 131  
 DB 334 GGGCTGCCACGACGCGCGACCCGCGCGTCGAGATCGCGCGCACTTCGCGCGCGCGCG 393  
 QY 132 GluGlnProValArgAsnLeuProValValGlyLysLeuProAspSerIleLysGly 151  
 DB 394 GAGAGCGCGCGGTCGACGAGTCCCGCTCTCCGCGCGCATCCCGCCCTTCATCAGCGGG 453  
 QY 152 ValTyrValArgAsnGlyAlaAsnProLeuHisGluProValThrGlyHisHisPhePhe 171  
 DB 454 GTCTACGCGCGCAACGGCGCCCAACCCCTGCTTCGACCCCGTCCGCGGGGACACCATCTTC 513  
 QY 172 AspGlyAspGlyMetValHisAlaValLysPheGluHisGlySerAla---SerTyrAla 190  
 DB 514 GACGCGACGCGATGTCGACGCGTCGCGATACGCAACGCGCGCGCGCGAGTCTCAGCGC 573  
 QY 191 CysArgPheThrGlnThrAsnArgPheValGlnGluArgGlnLeuGlyArgProValPhe 210  
 DB 574 TGCGCGCTTCAGGAGACCGCGCGCTGCGCAGGAGCGCGGATCGCGCGCGCGCGCTTTC 633  
 QY 211 ProLysAlaIleGlyGluLeuHisGlyHisThrGlyIleAlaArgLeuMetLeuPheTyr 230  
 DB 634 CCCAAGGCCATTGGCGAGCTGCACGGGCACTCCGGGATCGCGCGCTCGCCCTGTCTAC 693  
 QY 231 AlaArgAlaAlaGlyIleValAspProAlaHisGlyThrGlyValAlaAsnAlaGly 250



```
Db 694 GCGCGCGCGTGGCGCTCGTGGACCCCTCGCGCGGCACCGCGTGGCCAAACGCGGC 753
Qy 251 LeuValTyrPheAsnGlyArgLeuLeuAlaMetSerGluAspLeuProTyrGlnVal 270
Db 754 CTGCTACTTCAACGGCGCGCTGCTGCCATGTCGAGAGGACGACCTCCCTACACGTC 813
Qy 271 GlnIleThrProAsnGlyAspLeuLysThrValGlyArgPheAspPheAspGlyGlnLeu 290
Db 814 CGGTGGCGGACGAGCGGACCTCGAGACCTCGCGCGCTCGAGCTTCGAGCGGAGCTC 873
Qy 291 GluSerThrMetIleAlaHisProLysValAspProGluSerGlyGluLeuPheAlaLeu 310
Db 874 GGCTGGCGCATGATCGCGCACCCAAAGCTGGACCGGCCACCGGGAGCTCCACGCGTC 933
Qy 311 SerTyrAspValSerLysProTyrIleuLysTyrPheArgPheSerProAspGlyThr 330
Db 934 AGTACGACGTCATCAGAGCGCGTACTCCTCAAGTACTTCTACTTCAGGCGCGGACCC 993
Qy 331 LysSerProAspValGluIleGlnLeuAspGlnProThrMetMetHisAspPheAlaIle 350
Db 994 AAGTCGACGACGTGGAGATCCGCTGGAGCAGCCCGACGATGATCCAGACTTCGCCATC 1053
Qy 351 ThrGluAsnPheValValProAspGlnGlnValValPheLysLeuProGluMetIle 370
Db 1054 ACCGAGAACTTCGTGGTGTGCCCGCACCCAGGTGGTGTTCAGCTCCAGGAGTGTG 1113
Qy 371 ArgGlyGlySerProValValTyrAspLysAsnLysValAlaArgPheGlyIleLeuAsp 390
Db 1114 CGCGCGCGTGGCGCGTGTGTCGACAAAGGAGAGACGTCGCGGTTCGGCGTCTCC 1173
Qy 391 LysTyrAlaGluAspSerSerAsnIleLysTrpIleAspAlaProAspCysPhe 410
Db 1174 AAGCACGCGCGGACGCGTGGAGATCGCGTGGGTGGACGTCGCGACTTCCTGCTTC 1233
Qy 411 HisLeuTrpAsnAlaTrpGluGluProGluThrAspGluValValIleGlySerCys 430
Db 1234 CACCTGTGAACGCGTGGAGACGAGCGGCGGAGGAGTGGTGTGTCGCTCGTC 1293
Qy 431 MetThrProProAspSerIlePheAsnGlnSerAspGluAsnLeuLysSerValLeuSer 450
Db 1294 ATGACCCCGCGGACTCCTCATCTTCAACGAGTCCGACGAGCGCTGGAGAGTGTGCACC 1353
Qy 451 GluIleArgLeuAsnLeuLysThrGlyGluSerThrArgArgProIleIleSerAsnGlu 470
Db 1354 GAGATCCCGCTGGACCGCGGCGCGGCTCCAGCGCGCGCGCTCTCCCG---CCG 1410
Qy 471 AspGlnGlnValAsnLeuGluAlaGlyMetValAsnArgAsnMetLeuGlyArgLysThr 490
Db 1411 TCGCAGCAGGAGAACCTCGAGTGGCATGTGTGAACCGCAACCTGCTGGGCGCGAGAGC 1470
Qy 491 LysPheAlaTyrLeuAlaLeuAlaGluProTrpProLysValSerGlyPheAlaLysVal 510
Db 1471 CGGTACGCGTACCTCGCGTGGGAGCGGTGGCCCAAGAGAGTGGGCTTCGCCAAGGAG 1530
Qy 511 AspLeuThrThrGlyGluValLysLysHisLeuTyrGlyAspAsnArgTyrGlyGlyGlu 530
Db 1531 GACCTGTCCAGCGGGGAGCTCACCAGTTCGAGTACGCGGAGCGCGGTTCGGCGGCGAG 1590
Qy 531 ProLeuPheLeuProGlyGlyGlyGlyGlu-----GluAspGluGlyTyr 545
Db 1591 CCCTGCTTCGTTCCATGACCCCGCGCGCGCCACCCCGCGCGGAGGACGCGGTAC 1650
Qy 546 IleLeuCysPheValHisAspGluLysThrTrpLysSerGluLeuGlnIleValAsnAla 565
Db 1651 GTGCTCACTTCGTCCACGAGGAGCGCGCGACGTCGAGCTACTTGTGTCAATGCC 1710
Qy 566 ValSerLeuGluValGluAlaThrValLysLeuProSerArgValProTyrGlyPheHis 585
Db 1711 GCCGACATCCCGGTGGAGCCACGGTTCAGCTGCCGTCCCGCGTCCCTTCGGCTTCAC 1770
Qy 586 GlyThrPheIleGlyAlaAspLeuAlaLysGln 597
Db 1771 GGCACCTTCATCACGGCGGACGAGCTCGAGGCCAC 1806
```

## RESULT 6

AAD09398 standard; cDNA; 1734 BP.

AC AAD09398;

XX 10-SEP-2001 (first entry)

XX Arabidopsis thaliana neoxanthin cleavage enzyme, AtNCED5 cDNA.

DE Neoxanthin cleavage enzyme; AtNCED5; abscisic acid; ABA; herbicide;  
KW stress tolerance; transgenic plant; plant breeding; antisense-therapy;  
KW plant growth protectant; ss.

XX Arabidopsis thaliana.

PH Key Location/Qualifiers

FT CDS 1..1734

FT /\*tag= a

FT /product= "Arabidopsis thaliana AtNCED5 protein"

XX EP1116794-A2.

PD 18-JUL-2001.

XX 11-JAN-2001; 2001EP-0300218.

XX 13-JAN-2000; 2000JP-0010056.

PR 11-JAN-2001; 2001JP-0003476.

XX (RIKE ) RIKEN KK.

XX Iuchi S, Kobayashi M, Shinozaki K;

XX WPI; 2001-400081/43.

DR P-PSDB; AAE04786.

XX A DNA encoding a protein with a neoxanthin cleavage activity for  
producing transgenic plants with improved or decreased stress tolerance

XX Claim 3; Page 46-49; 101pp; English.

XX The invention relates to neoxanthin cleavage enzymes and their  
corresponding cDNA molecules. Neoxanthin cleavage enzyme plays a key  
role in endogenous abscisic acid (ABA) biosynthesis under drought stress.  
XX Neoxanthin cleavage enzyme is used for improving stress tolerance in a  
plant when expressed in a plant cell. The invention also relates to  
XX methods for increasing or decreasing stress tolerance in a plant by  
introducing the DNA into the plant, and a transgenic plant into which a  
neoxanthin cleavage enzyme is introduced. The improvement of stress  
XX tolerance in plants is useful, for example in plant breeding. Neoxanthin  
cleavage enzyme genes are useful for producing transgenic plants. An arid  
XX land can be improved by growing transformant weed for several years and  
then removing the weed by specifically lowering stress tolerance in the  
weed by inducing an inducible promoter. The present cDNA sequence encodes  
XX Arabidopsis thaliana neoxanthin cleavage enzyme, AtNCED5 protein.  
The AtNCED5 cDNA is obtained from an Arabidopsis plant-derived cDNA  
XX library using a cDNA of the CPRD65 (CowPea Responsive to Dehydration)  
gene isolated from cowpea plant as a probe.

SQ Sequence 1734 BP; 464 A; 390 C; 441 G; 439 T; 0 other;

Alignment Scores:

Pred. No.:	1..2e-152	Length:	1734
Score:	1663.50	Matches:	323
Percent Similarity:	72.56%	Conservative:	100
Best Local Similarity:	55.40%	Mismatches:	137
Query Match:	52.81%	Indels:	23
DB:	22	Gaps:	9

US-09-758-269-6 (1-599) x AAD09398 (1-1734)



```
QY 24 LeuSerSerSerGlnSerSerLeuSerTyrcysSerSerLeuProMetAlaSerArg 43
Db 25 CTTCTTCCGACGAGACTTCTCTCGTCTCTCTTACTTCTCCACACCCCAAAATGCAAT 84
QY 44 ValThrArgLysLeuAsnValSerSer---AlaLeuHisThrProProAlaLeuHisPhe 62
Db 85 ATTCTCGAGCAATCTCTTAACCTTTCAGATACCGACATCTCTCTGATCTCTCTTCT 144
QY 63 ProLysGlnSerSerAsnSerProAlaLeuValLysProLysAlaLysGluSerAsn 82
Db 145 CCGTTCCGCTCACCG-----GTTAAGCTCAACCAACCAAGTATCCAAAC 186
QY 83 ThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaLeuAspAlaGlu--- 101
Db 187 -----TTAAACCTTCTTTCAGAGCTAGCGGCTTACGATCTCGACAAGATTGAGTCC 237
QY 102 GlyPheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSerVal 121
Db 238 TCTATCGTTATTCTCTATGAGCAGCAATCGCCGCTTCTTAAACCGACCGCCGCGGTT 297
QY 122 GlnIleAlaGlyAsnPheAlaProValAsnGluGlnProValArgAsnLeuProVal 141
Db 298 CAATTATCAGGTAACTTCTCTCGGTTAATGATGCTCGGTTTCAGAACCGTTTAGAAGTG 357
QY 142 ValGlyLysLeuProAspSerIleLysGlyValTyrrValArgAsnGlyAlaAsnProLeu 161
Db 358 GTTGTGTCAGATCTCTTCTGTCTAAAGAGGATTACATCCGTAAACCGGTCAAAACCTATG 417
QY 162 HisGluProValThrGlyHisPheAspGlyAspGlyMetValHisAlaValLys 181
Db 418 TTTCCGCGCTTACCGCGGACATCATTTATTTGACGCTGACGGAATGATTCACCGCTTAGT 477
QY 182 -----PheGluHisGlySerAlaSerTyrrAlaCysArgPheThrGlnThrAsnArgPhe 199
Db 478 ATCGGTTTGTATAAC---CAGGTTAGTACAGCTGCCGTTACACTTAAACCAACCGGCTT 534
QY 200 ValGlnGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyLysLeuHisGly 219
Db 535 GTTCAAGAAACCGCGCTTGGACGATCGGTTTCCCTTAAACCAATCGCGAGCTTCACGCG 594
QY 220 HisThrGlyIleAlaArgLeuMetLeuPheTyrrAlaArgAlaAlaGlyIleValAsp 239
Db 595 CAFTCCGCTAGCTAGCTCGCTCTCTCACGCTCGAGCTGCGGATCGGTTCTAGTGGAC 654
QY 240 ProAlaHisGlyThrGlyValAlaAsnAlaGlyLeuValTyrrPheAsnGlnValArgLeu 259
Db 655 GGGACACGTGGCATCGGCTAGCTAACGCCGCTGCTGTTCTTTTAAACGCGAGGTATTA 714
QY 260 AlaMetSerGluAspLeuProTyrrGlnValGlnIleThrProAsnGlnValAspLeuLys 279
Db 715 GCCATGTTCAGAAAGATGATCTCTTACCAAGTGAAGATCGACGCTCAAGGAGATCTTGAG 774
QY 280 ThrValGlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLys 299
Db 775 ACGATTCGACCGGTTCCGATTCGATACCCAGATGACTCTCTAGTATAGCGCATCTTAAG 834
QY 300 ValAspProGluSerGlyGluLeuPheAlaLeuSerTyrrAspValValSerLysProTyrr 319
Db 835 GTGGACGCGCACACAGGAGATCTTCATACACTGACTACCACTTTTGAAGAACTCAT 894
QY 320 LeuLysTyrrPheArgPheSerProAspGlyThrLysSerProAspValGluIleGlnLeu 339
Db 895 CTCAGGTATCTTAAATTCACACGCTCGCGGAAAGACACGCTGAGTGGAGATCACGCTC 954
QY 340 AspGlnProThrMetMetHisAspPheAlaIleThrGluAsnPheValValProAsp 359
Db 955 CCGTAACCAACGATGATCATGATTCGCGATAACCGAGAAATTTGTCTGTTATACCGGAT 1014
QY 360 GlnGlnValValPheLeuLeuProGluMetIleArgGlyGlySerProValValTyrrAsp 379
Db 1015 CAGCAATGGTATTCAAAATTAATCCGAAATGATTCGGGGCGGTACCCGTTATCTACGTT 1074
```

```
QY 380 LysAsnLysValAlaArgPheGlyIleLeuAspLysTyrrAlaGluAspSerSerAsnIle 399
Db 1075 AAAGAAAAAATGCGGAGATTTTGGATTTTGTCAAAAGCAGGATCTGACCGGTCGGATATA 1134
QY 400 LysTrrPilleAspAlaProAspCysPheCysPheHisLeuTrpAsnAlaTrpGluPro 419
Db 1135 AATTGGGTTGATGTACCGGATTTGTTTCTTTCATCTATGGAATGCGTGGGAAGAG--- 1191
QY 420 GluThrAspGlu-----ValValValIleGlySerCysMetThrProProAsp 435
Db 1192 AGAACCGAGAGGAGAGACCGGATTCGTCGTAATCGGTCATGTATGAGCCACCCGAC 1251
QY 436 SerIlePheAsnGluSerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsn 455
Db 1252 ACGATCTTTAGTGAATCAGGAGACCAACCCGGTGAATTAAGTGAGATCCGGTTAAAC 1311
QY 456 LeuLysThrGlyGluSerThrArgArgProIleIleSerAsnGluAspGlnGlnValAsn 475
Db 1312 ATGCGTCAAAAGAAATCGAACCGTAAAGGTTATCGTAACCCGGA-----GTGAAT 1359
QY 476 LeuGluAlaGlyMetValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrrLeu 495
Db 1360 TTAGAACGGGTCACATAAACCGTAGTTACGTGGCCCGGAAAGCCAGTTCGTTTACATA 1419
QY 496 AlaLeuAlaGluProTrpProLysValSerGlyPheAlaLysValAspLeuThrThrGly 515
Db 1420 GCAATAGCCGATCTCTTGGCCCAATGTCAGTGGCATTGCGAAGGTAGATATACAAAACGCG 1479
QY 516 GluValLysLysHisLeuTyrrGlyAspAsnArgTyrrGlyGluProLeuPheLeuPro 535
Db 1480 ACCGTTTCAGAGTTTAAATTACGACCGGCGGTTTCGTTGGAGAACCGTCTTTGTACCG 1539
QY 536 GlyGluGlyGlyGluGluAspGluGlyTyrrIleLeuCysPheValHisAspGluLysThr 555
Db 1540 GAGGAGAGAGAGAGAGAGAGACAAAGTTATGTATGGGTTTGTGAGAGACGAGAGAAA 1599
QY 556 TrpLysSerGluLeuGlnIleValAsnAlaValSerLeuGluValGluAlaThrValLys 575
Db 1600 GACGAGTCGAGGTTTGTGTTGTCGACGCGAGGATATGAAGCAAGTCGCGCGGTGCGC 1659
QY 576 LeuProSerArgValProTyrrGlyPheHisGlyThrPheIleGlyAlaAspAspLeuAla 595
Db 1660 TTGCGGAGAGGGTACCTTATGGTTTTCATGGAACGTTTCGTGAGCGAGAAATCAGTTGAAG 1719
QY 596 LysGlnVal 598
Db 1720 GAACAAGTT 1728
RESULT 7
AAC42989
ID AAC42989 standard; DNA; 1788 BP.
XX AC AAC42989;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 37611.
XX KW Hybridisation assay; genetic mapping; gene expression control;
XX KW protein identification; signal transduction pathway;
XX KW metabolic pathway; promoter; terminator; termination sequence; ss.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
```

PR 23-MAR-1999; 99US-0135788.  
PR 25-MAR-1999; 99US-0136264.  
PR 29-MAR-1999; 99US-0136786.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130444.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135622.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137503.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139453.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0143290.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143544.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.

PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

## Alignment Scores:

Pred. No.: 1,21e-84 Length: 1788  
 Score: 969.00 Matches: 214  
 Percent Similarity: 55.31% Conservative: 114  
 Best Local Similarity: 36.09% Mismatches: 221  
 Query Match: 30.76% Indels: 44  
 DB: 21 Gaps: 13

US-09-758-269-6 (1-599) x AAC42989 (1-1788)

QY 32 LeuSerTyrCysSerSerLeuProMetAlaSerArgValThrArgLysLeuAsnValSer 51  
 DB ||| : : : : : ||| : : : : :  
 56 CTCTCTCTCGCGCGATCTTCTCTC-----CTACTC 88  
 QY 52 SerAlaLeuHisThrProAlaLeuHisPheProLysGlnSer-----SerAsn 68  
 DB ||| : : : : : ||| : : : : :  
 89 TCTCTCGATCACTCCGCGCTGTCGAAGACGTTCTCCAATCAAAACCCCAACGAC 148  
 QY 69 SerProAlaIleValLysProLysAlaLysGluSerAsnThrLysGlnMetAsnLe 88  
 DB ||| : : : : : ||| : : : : :  
 149 ACAATGATCGTGTAAACAAACCAAACTCCACCAACCAACCAACCACTTAGTCT 208  
 QY 88 uPheGlnArgAlaAlaAlaAlaLeuAspAlaAlaGluGlyPheLeuValSerHisG 108  
 DB ||| : : : : : ||| : : : : :  
 209 CATCACCAACCAACCACTCCGACGAGAAATGACTCTCGCAACGCTCTCTTCCACC 268  
 QY 108 LuLysLeu-----HisProLeuProLysThrAlaAspProSerValG 122  
 DB ||| : : : : : ||| : : : : :  
 269 AAGATGTAATCAACAGCTTCATCGATCCACCTTCACGCTTCCTTCGTTGATCAAA 328  
 QY 122 InIleAlaGlyAsnPheAlaProValAsnGluGlnProValArgAsnLeuProVal 142  
 DB ||| : : : : : ||| : : : : :  
 329 TCCTCTCTGTAACCTTCGCTCTGCTCTCGACGAGCTTCTCCAACAGACTGTGAATCA 388  
 QY 142 al---GlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyAlaAsnProL 161  
 DB ||| : : : : : ||| : : : : :  
 389 TCCACGGCACTCTTCACCTGTCACTTAACGGCGCTTACATCGTAACGGTCCAAATCCAC 448  
 QY 161 euHisGluProValThrGlyHisHisPhePheAspGlyAspGlyMetValHisAlaValL 181  
 DB ||| : : : : : ||| : : : : :  
 449 AGTTTCTCCCTCGTGGTCTTACCATCTCTTCGACGGCGACGGTATGCTTACGCGCAT 508

QY 181 yePheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheValG 201  
 DB ||| : : : : : ||| : : : : :  
 509 AAATCCACAACCGTAAAGCCACTCTCTGTAGCAGATACGTCAGACTTATAAATACAACG 568  
 QY 201 InGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGlyHisT 221  
 DB ||| : : : : : ||| : : : : :  
 569 TCGAGAAACAAACCGGAGCTCCGTTATGCCTAACGCTGTTTCCGGATTCAACGGGTAA 628  
 QY 221 hr---GlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaAlaGlyIleValAsp 240  
 DB ||| : : : : : ||| : : : : :  
 629 CGGCTCAGTAGCTCGTGAGCTTTAACGGCAGCTTAGGTTTAAACCGGACAGTATAATC 688  
 QY 240 roAlaHisGlyThrGlyValAlaAsnAlaGlyLeuValTyrPheAsnGlyArgLeuLeuA 260  
 DB ||| : : : : : ||| : : : : :  
 689 CGGTTAACGGCATTCGTTTAGCTAATACAAGTCTAGCTTCTTCAGTAACCGCTCTCTTG 748  
 QY 260 laMetSerGluAspAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysT 280  
 DB ||| : : : : : ||| : : : : :  
 749 CTTTAGGTGAATCTCATTTACCTACGCGCTCCGATTAAACCAATCAGGAGATATTGAAA 808  
 QY 280 hrValGlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysV 300  
 DB ||| : : : : : ||| : : : : :  
 809 CGATCGGACGGTACGATTTTCGACGGGAAATTAGCGATGATGATGACAGCTCATCTTAAA 868  
 QY 300 alAspProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValSerLysProTyrL 320  
 DB ||| : : : : : ||| : : : : :  
 869 CCGATCCAATAACCGGAGAACTTTCGCTTCCGGTACGGTCCGGT---CCACCGGTTT 925  
 QY 320 euLysTyrPheArgPheSerProAspGlyThrLysSerProAspValGluLle---GlnL 339  
 DB ||| : : : : : ||| : : : : :  
 926 TAACATATTTCCGGTTTGAATTCGCGCGGAAAAACAAAGAGACGTTCCGATATTCTCGA 985  
 QY 339 euAspGlnProThrMetMetHisAspPheAlaIleThrGluAsnPheValValProA 359  
 DB ||| : : : : : ||| : : : : :  
 986 TGACGCTCCGTCGTTCTCCATGACTTCGCGATCAGGAAACGTCACCGGATTTTCGCGAG 1045  
 QY 359 spGlnGlnValValPheLys-----LeuProGluMetIleArgGlyGlySerProV 376  
 DB ||| : : : : : ||| : : : : :  
 1046 AGATTCAGCTTGGCATGAGGATGAACATGTTGGATTGTTCTCGAAGTGCTTCTCCG 1105  
 QY 376 alValTyrAspLysAsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspS 396  
 DB ||| : : : : : ||| : : : : :  
 1106 TTGTTACTGATAACCGGAAAACTCCAGAGCTTGGAGTGAATTCCTAAGTACGCGGAGATG 1165  
 QY 396 erSerAsnIleLysTrpIleAspAlaProAspCysPheCysPheHisLeuTrpAsnAlaT 416  
 DB ||| : : : : : ||| : : : : :  
 1166 AGTCGAGATGAATGGTTCCGAAGTTCCTCGATTCAATATCATTCACGCTATTATGCTT 1225  
 QY 416 rpGluGluProGluThrAspGluValValIleGlySerCysMetThrProProAspS 436  
 DB ||| : : : : : ||| : : : : :  
 1226 GGGATGAAGATGATGGAACACGCTCGTTTGTATTCACCGAATATTATGTCGATTGAAC 1285  
 QY 436 erIlePheAsnGluSerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsnL 456  
 DB ||| : : : : : ||| : : : : :  
 1286 ATACTTTAGAGAGGATGAT---CTGTTTCATGCTTTCGTCGAGAGGTCGAAGATCGATC 1342  
 QY 456 euLysThrGlyGluSerThrArgArgProIleIleSerAsnGluAspGlnGlnValAsnL 476  
 DB ||| : : : : : ||| : : : : :  
 1343 TCGTCACCGGATTTGTGAGACGTCATCGATCTCAGCGAGG-----AATC 1387  
 QY 476 euGluAlaGlyMetValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeuA 496  
 DB ||| : : : : : ||| : : : : :  
 1388 TCGATTCGCTGTGATTAATCCGGCGTTTTCGCGGAGATGTAGCAGGTACGTTTACGCGG 1447  
 QY 496 laLeuAlaGluProTrpProLysValSerGlyPheAlaLysValAspLeuThrThrGlyG 516  
 DB ||| : : : : : ||| : : : : :  
 1448 CGATTGGAGATCCGATGCGGAAGATCTCCGGTGGTGGTGAAGCTTGATGTCTAAAGGAG 1507  
 QY 516 lu-----ValLysLysHisLeuTyrGlyAspAsnArgTyrGlyGlyGluP 531  
 DB ||| : : : : : ||| : : : : :  
 1508 ATCGGATGATTTGTACGCTGCCCGTAGAATGTAGCTTACGGTGTGTTACGCGGAGAAC 1567



```
Db 869 CCGATCCAAATACCGGAGAACTTTTCGTTTCGGTACCGTCCGGTT---CCACCGTTT 925
Qy 320 euLysTyrPheArgPheSerProAspGlyThrLysSerProAspValGluIle--GlnL 339
Db 926 TAACATATTTCGGTTTGGATTCGCCGCGGAAAAACAAAGAGACGTTCCGATATTCCGA 985
Qy 339 euAspGlnProThrMetMetHisAspPheAlaIleThrGluAsnPheValValProA 359
Db 986 TGACGCTCCGTTCTCCATGCTTCGATGCTTCGATGCTTCGATGCTTCGATGCTTCG 1045
Qy 359 spGlnGlnValValPheLys-----LeuProGluMetIleArgGlyGlySerProV 376
Db 1046 AGATTCCGTTGGCATGAGATGAACATGTTGGATTGGTTTCTCGAAGTGGTTCTCCG 1105
Qy 376 alValTyrAspLysAsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAsp 396
Db 1106 TTGGTACTGATAACGGAAGAACTCCAAGGCTTGGAGTATTCTTAAGTACGCGGAGATG 1165
Qy 396 erSerAsnIleLysTrpIleAspAlaProAspCysPheHisLeuTrpAsnAlaT 416
Db 1166 AGTCGAGATGAATGGTTTCGAAGTTCCTGGATTCAATATCATTCACGCTATTATGCTT 1225
Qy 416 rpGluGluProGluThrAspGluValValIleGlySerCysMetThrProAspAs 436
Db 1226 GGGATGAGATGATGAACAGCGTCTGTTTATTCACCGCATATTTATGTCATTGAAC 1285
Qy 436 erIlePheAsnGluSerAspGluAsnLysSerValLeuSerGluIleArgLeuAsnL 456
Db 1286 ATACTTTAGAGAGGATGGAT---CTGGTTTCATGCTTGGTGGAAGGTGAGATCGATC 1342
Qy 456 euLysThrGlyGluSerThrArgArgProIleSerAsnGluAspGlnGlnValAsnL 476
Db 1343 TCGTCACCGGGATTGTGACAGCTCATCCGATCTCAGCGAGG-----AATC 1387
Qy 476 euGluAlaGlyMetValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeuA 496
Db 1388 TCGATTTCGCTGTGATTAATCCGCGCTTCTCGGAGATGTAGCAGGTACGTTTACGCGG 1447
Qy 496 laLeuAlaGluProTrpProLysValSerGlyPheAlaLysValAspLeuThrThrGly 516
Db 1448 CGATTGGAGATCGATGCCGAGATCTCCGGTGTGGTGAAGCTTATGTCGTAAGGAG 1507
Qy 516 lu-----ValLysLysHisLeuTyrGlyAspAsnArgTyrGlyGlyGluP 531
Db 1508 ATCGGATGATTGTACGTTGGCCCGTAGAATGTACGTTTCAGGTTTACGCGCGAGAAC 1567
Qy 531 roLeuPheLeuProGlyGluGlyGly-----GluGluAspGluGlyTyrIleL 547
Db 1568 CGTTTTTCGTAGTACGATAGGATCTCGGTAATCCGAGCGGAGGAGGATGATGTTATG 1627
Qy 547 euCysPheValHisAspGluLysThrTrpLysSerGluLeuGlnIleValAsnAlaVal 567
Db 1628 TCAGTATGTTTCAGATGAAGTACTGGAGATCGAAGTTTCTGGTATGACGCTAAT 1687
Qy 567 er-----LeuGluValGluAlaThrValLysLeuProSerArgValProTyrGlyPheH 585
Db 1688 CGCCGAGGCTTGAATTCGCGCGCGGTGAGGTTCCGCGAAGGTTCCGTACGATGCC 1747
Qy 585 igGlyThrPheIleGlyAlaAspAspLeuAlaLys 596
Db 1748 ATGGGTTATTGTCAGGAAGAAAGTACCTTAATAAG 1782
RESULT 9
ID ABZ13639 standard; DNA; 1788 BP.
XX AC ABZ13639;
XX DT 21-JAN-2003 (first entry)
XX DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1444.
XX
```

```
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO200216655-A2.
XX
PD 28-FEB-2002.
XX
XX 24-AUG-2001; 2001WO-US26685.
XX
XX 24-AUG-2000; 2000US-227866P.
XX
XX 26-JAN-2001; 2001US-264847P.
XX
XX 22-JUN-2001; 2001US-30011P.
XX
XX (SCRI ) SCRIPPS RES INST.
XX
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
XX Harper JF, Kreps J, Wang X, Zhu T;
XX
XX WPI; 2002-304127/34.
XX
XX Identifying a stress condition to which a plant cell has been exposed
XX and producing plants with increased tolerance to these abiotic stresses
XX
XX Claim 144; SEQ ID NO 1444; 577pp + Sequence Listing; English.
XX
XX The invention relates to identifying a stress condition to which a plant
XX cell has been exposed, comprising:
XX (a) contacting nucleic acid representative of expressed polynucleotides
XX in the plant cell with an array or probes representative of the plant
XX cell genome; and
XX (b) detecting a profile of expressed polynucleotides in the plant cell
XX characteristic of a stress response. The method is useful in the
XX production of transgenic plants, cells and seeds and in producing plants
XX with increased tolerance to abiotic stress. The present sequence is that
XX of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
XX in methods of the invention.
XX Note: The sequence data for this patent is not represented in the printed
XX specification but is based on sequence information supplied to Derwent by
XX the European Patent Office.
XX
XX Sequence 1788 BP; 444 A; 444 C; 418 G; 482 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 1.21e-84 Length: 1788
XX Score: 969.00 Matches: 214
XX Percent Similarity: 55.31% Conservative: 114
XX Best Local Similarity: 36.09% Mismatches: 221
XX Query Match: 30.76% Indels: 44
XX DB: 24 Gaps: 13
XX
XX US-09-758-269-6 (1-599) x ABZ13639 (1-1788)
Qy 32 LeuSerTyrCysSerSerLeuProMetAlaSerArgValThrArgLysLeuAsnValSer 51
Db 56 CTCTTCTTCGCGCGCATCTTCTCTCTC-----CTACTC 88
Qy 52 SerAlaLeuHisThrProAlaLeuHisPheProLysGlnSer-----SerAsn 68
Db 89 TCCTCCGTTATCAACTCCGCGCTCGTGAAGAACGTTCTTCCCAATCAAAACCAACGACGACA 148
Qy 69 SerProAlaIleValValLysProLysAlaLysGluSerAsnThrLysGlnMetAsnLe 88
Db 149 ACAATGATCGTGTAAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 208
Qy 88 uPheGlnArgAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 108
Db 209 CATCACCACCCGAACTCCGACGAGAAATGACTCTCTGCAACAGCTCTCTTACCACCGTCG 268
Qy 108 luLysLeu-----HisProLeuProLysThrAlaAspProSerValG 122
Db 269 AAGATGTAATCAACACACGTTTCATCGATCCACCTTCACGTCCTTCCGTTGATCCAAACATG 328
```

```
Qy 122 lnIleAlaGlyAsnPheAlaProValAsnGluGlnProValArgAsnLeuProValV 142
Db 329 TCCTCTCTGATAAAGTTCCTGCTCGGACGAGCTTCCTCCACAGACGTGTAATCA 388
Qy 142 al--GlyLeuLeuProAspSerIleLeuValValArgAsnGlyAlaAsnProL 161
Db 389 TCCAGGACACTTCCACTGTCATTAACGGGCTTACATCCGTAAACGGTCCAAATCCAC 448
Qy 161 euHisGluProValThrGlyHisHisPheAspGlyAspGlyMetValHisIleValL 181
Db 449 AGTTTCTCCTCGTGTCTTACCATCTCTTCGACGGGACGGTATGCTTCACGCCATAA 508
Qy 181 yspPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheValG 201
Db 509 AAATCCACACGGTAAACCCACTCTCTGTACACATACGTCAAGACTTATTAATACACG 568
Qy 201 lngUArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGlyHisT 221
Db 569 TCGAGAAACAACCGGAGCTCGGTATGCTTAACGTTTTCGGATTCAACGGTGTA 628
Qy 221 hr--GlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaGlyIleValAspP 240
Db 629 CGGCTCAGTAGCTCGTGAGCTTTAACGGCAGCTAGGGTTTAAACCGCAGATATAATC 688
Qy 240 roAlaHisGlyThrGlyValAlaAsnAlaGlyLeuValTyrPheAsnGlyArgLeuLeuA 260
Db 689 CGGTTAACGGCATGTTGGTTAGCTAATACAGTCTAGCTTCTTCAGTAAACCGTCTCTTTG 748
Qy 260 laMetSerGluAspAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysT 280
Db 749 CTTTAGTGATCTGATTTTACCTACGCGCTCCGATTAACCGAATCAGAGATATTGAA 808
Qy 280 hrValGlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysV 300
Db 809 CGATCGGACGGTACGATTCGACGGAAATTAGCGATGATGATGATGATGATGATGATGAT 868
Qy 300 alAspProGluSerGlyLeuLeuPheAlaLeuSerTyrAspValValSerLysProTyrL 320
Db 869 CCGATCCAAATACCGGAGAAATCTTCGCTTTCCGGTACGGTCCGGT---CCACCGTTT 925
Qy 320 euLysTyrPheArgPheSerProAspGlyThrLysSerProAspValGluIle---GlnL 339
Db 926 TAACATATTTCCGTTTTCGATTCGCGCGGAAAAACAAGAGACGTTCCGATATCTCGA 985
Qy 339 euAspGlnProThrMetHisAspPheAlaIleThrGluAsnPheValValProA 359
Db 986 TGACGTCTCCGTCGTTTCTCCATGACTTCGGATCACGAAACGTCACGCGATTTTCGCAG 1045
Qy 359 spGlnGlnValValPheLys-----LeuProGluMetIleArgGlyGlySerProV 376
Db 1046 AGATTACAGCTTGGCATGAGGATGAACATGTTGGATTTTCGTTCTCGAAGGTGGTTCTCCGG 1105
Qy 376 alValTyrAspLysAsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspS 396
Db 1106 TTGGTACTGATTAACCGAAATAAATCCAGGCTTGGAGTGATTCCTTAAGTACCGCGGAGATG 1165
Qy 396 erSerAsnIleLysTrpIleAspAlaProAspCysPheCysPheHisLeuTrpAsnAlaT 416
Db 1166 AGTCGGAGATGAAATGGTTTCGAAGTTCCTGGATTCAATATCATTCACGCTATTATGCTT 1225
Qy 416 rpGluGluProGluThrAspGluValValValIleGlySerCysMetThrProProAspS 436
Db 1226 GGGATGAAGATGATGGAACACGCGCTGTTTGGATTGCACCGAATATTATGCGATTGAAC 1285
Qy 436 erIlePheAsnGluSerGluAsnLeuLysSerValLeuSerGluIleArgLeuAsnL 456
Db 1286 ATACTTTAGAGAGGATGAT---CTGGTTCATGCTTTGGTGAGAGGAGTGAGATCGATC 1342
Qy 456 euLysThrGlyLeuSerThrArgArgProIleIleSerAsnGluAspGlnGlnValAsnL 476
Db 1343 TCGTCACCGGGATTGTGAGACGTCATCCGATCTCAGCGAGG-----AATC 1387
```

```
Qy 476 euGluAlaGlyMetValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeuA 496
Db 1388 TCGATTTTCGTGTGATTAAATCCGGCGTTTCTCGGAGATGTAGCAGGTACGTTTACCGGG 1447
Qy 496 laLeuAlaGluProTrpProLysValSerGlyPheAlaLysValAspLeuThrThrGlyG 516
Db 1448 CGATTGGAGATCCGATCCGAAGATCTCCGTTGGTGAAGCTTGTATGTCTAAAGGAG 1507
Qy 516 lu-----ValLysLysHisLeuTyrGlyAspAsnArgTyrGlyGlyGluP 531
Db 1508 ATCGGGATGATTGACCGTGGCGTAGAATGTACGGTTTTCAGGTGTTTACCGCGGAGAAC 1567
Qy 531 roLeuPheLeuProGlyGlyGly-----GluGluAspGluGlyTyrIleL 547
Db 1568 CGTTTTCGTAGTAGGATCTCTGTAATCCGAGGCGGAGGAGATGATGTTATGTGG 1627
Qy 547 euCysPheValHisAspGluLysThrTrpLysSerGluLeuGlnIleValAsnAlaValS 567
Db 1628 TGACGTATGTTTACAGTGAAGTGAAGTGAAGTTCGAGTTCGAGTTCGAGTGAAGT 1687
Qy 567 er-----LeuGluValGluAlaThrValLysLeuProSerArgValProTyrGlyPheH 585
Db 1688 CGCCGAGCTTGAATCGTCCCGCTGAGGTTGCCGCGAAGGTTCCCGTACGATGCC 1747
Qy 585 isGlyThrPheIleGlyAlaAspAspLeuAlaLys 596
Db 1748 ATGGTTATTGTCAAGAAAGTGCCTTAATAAG 1782
RESULT 10
AAF77206
ID AAF77206 standard; cDNA; 1950 BP.
XX
AC AAF77206;
XX
DT 16-MAY-2001 (first entry)
XX
DE cDNA encoding sunflower neoxanthin cleavage enzyme (NCE).
XX
KW Defence-related signalling gene; sunflower; neoxanthin cleavage enzyme;
KW NCE; amino acid permease; AAP; glutamic acid rich protein; GRP;
KW pathogen resistance; abscisic acid metabolism; ss.
XX
OS Helianthus annuus.
XX
FH Key Location/Qualifiers
FT CDS 1..1632
FT /tag= a
FT /product= "NCE"
FT /note= "Neoxanthin cleavage enzyme"
FT /partial
XX
XX WO200112801-A2.
XX
XX 22-FEB-2001.
XX
XX 17-AUG-2000; 2000WO-US22961.
XX
XX 18-AUG-1999; 99US-0149656.
XX 23-MAY-2000; 2000US-0206405.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX (CURA-) CURAGEN CORP.
XX
XX Bidney DL, Craeta OR, Hu X, Lu G;
XX
XX WPI; 2001-211215/21.
XX P-PSDB; AAB72303.
XX
XX Novel isolated defence-related signalling gene isolated from sunflower
XX encoding neoxanthin cleavage enzyme, amino acid permease or glutamic
XX acid-rich protein useful for increasing resistance of plant to a
XX pathogen -
XX
```

Claim 1; Page 94-97; 135pp; English.

PS This invention relates to defence-related signalling genes isolated from  
 CC the sunflower (*Helianthus annuus*). The genes encode a neoxanthin cleavage  
 CC enzyme (NCE), an amino acid permease (AAP) and a glutamic acid rich  
 CC protein (GRP). The signalling gene is useful for increasing the  
 CC resistance of a plant to a pathogen such as fungus, virus, bacterium,  
 CC nematode or insect (e.g. European corn borer), preferably  
 CC *Sclerotinia* spp., *Phoma* spp., or *Phomopsis* spp., by stably incorporating a  
 CC construct containing the gene into the genome of the plant. The gene is  
 CC useful for regulating gene expression in a plant, in response to a  
 CC stimulus such as infection with a pathogen, damage from a pathogen,  
 CC hydrogen peroxide, jasmonic acid, methyl jasmonate, salicylic acid,  
 CC oxalic acid or expression of a gene encoding oxalic acid oxidase. The  
 CC genes are also useful for stem-preferred regulation of gene expression in  
 CC a plant. The genes are useful in agriculture, particularly in the  
 CC breeding of crop plants with improved agronomic traits, for modifying  
 CC abscisic acid (ABA) metabolism and for modifying amino acid transport and  
 CC content in plants. The present sequence represents cDNA encoding the  
 CC sunflower neoxanthin cleavage enzyme (NCE).

XX SQ Sequence 1950 BP; 535 A; 456 C; 453 G; 506 T; 0 other;

#### Alignment Scores:

Pred. No.:	9-37e-84	Length:	1950
Score:	960.50	Matches:	209
Percent Similarity:	53.9%	Conservative:	109
Best Local Similarity:	35.4%	Mismatches:	214
Query Match:	30.4%	Indels:	57
DB:	22	Gaps:	13

US-09-758-269-6 (1-599) x AAF77206 (1-1950)

QY	51	SerSerAlaLeuHisThrProProAlaLeuHisPheProLysGlnSerSerAsnSerPro	70
DB	49	TCAACTGTTCCACCCACCACCA-----TCATCGTCTCACCG	87
QY	71	AlaIleValVal-----LysProLysAlaLysGluSer	81
DB	88	CCATCTGGCGCTAAGAGTCTTTTTCAGTCAGGACTGAAGAAACACCAACACCGCTCACCACT	147
QY	82	-----AsnThrLysGln-----MetAsnLeuPheGln	90
DB	148	ACCGCCACCCACAAAGAGCCTAGTCGGACCGAAACCAACCCCTGAATATTAGNAAA	207
QY	91	ArgAlaAlaAlaAlaLeuAspAla-----AlaGluGly	102
DB	208	CGTGAAGCTTCAGTCGGTGGTGCAGTCTTTACCGGCCACCACCTTTTAATGCTTTTGCAT	267
QY	103	PheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSerValGln	122
DB	268	AGCATCATTAATAACTTTATTGATCCACCGCGGAGGGTCTCAGTTGATCCAAACACCGTT	327
QY	123	IleAlaGlyAsnPheAlaProValAsnGlnGlnProValArgArgAsnLeuProValVal	142
DB	328	TGTCTGATAACTTTTCACCGGTGACGAACCTCCCTCCGACTGCTGAAAGTCATCGAG	387
QY	143	GlyLysLeuProAspSerIleLysGlyValThrValArgAsnGlyAlaAsnProLeuHis	162
DB	388	GGCACATGCCAAGTTCCTGACCGTCTTACTTCCGTAATGTCGCAACCCGCAATTC	447
QY	163	GluProValThrGlyHisPhePheAspGlyAspGlyMetValHisAlaValLysPhe	182
DB	448	CTTCGCGGAGGACCCCTACACCTCTTCGATGSCGATGCGTCTCCATGCTATTTCGTATC	507
QY	183	GluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheValGlnGlu	202
DB	508	TCCATGGAAAGCTTCGTTATGTAGTCGATCATCAAAACATACAAATATTTCATAGAG	567
QY	203	ArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGlyHisThr---	221
DB	568	AAAGAAAGCGGGATTCCCATTTATCCAAACGTGTTTTCAGGGTTTAAATGGTGTGACTGCC	627

QY	222	GlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaAlaGlyIleValAspProAla	241
DB	628	TCTGCAGCTCGCATGGCAGTCACTCCGCGCGGATTTTGGCTGGACAATTTGACCCACA	687
QY	242	HisGlyThrGlyValAlaAsnAlaGlyLeuValTyrPheAsnGlyArgLeuLeuAlaMet	261
DB	688	AAAGGTATTGGTCTAGCCAATACCACTGGCCCTTTTGGCAACAGACTTTTGTGCTCG	747
QY	262	SerGluAspAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysThrVal	281
DB	748	GGAGAGTCGGATCTCCCATATGCCGTCAACTAGCGCCCGCGGTGACATAGTACCCGTC	807
QY	282	GlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysValAsp	301
DB	808	GGACGTGAGGACTTCGACGGCAAACTATTCATGACATACCGCTCAGCCAAAATTCGAT	867
QY	302	ProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValSerLysProTyrLeuLys	321
DB	868	CCAGTAACGAAAGAGCTTTTGGCTTTCTGTTACGGTCCAGTC---CCCCCTTTCTTAACC	924
QY	322	TyrPheArgPheSerProAspGlyThrLysSerProAspValGluIle---GlnLeuAsp	340
DB	925	TTTTTCCGTTCAACGAAACCGAGAAAACAAGCCGATGTCCCGATCTTCTCATATGACA	984
QY	341	GlnProThrMetHisAspPheAlaIleThrGluAsnPheValValValProAspGln	360
DB	985	AGCCCGTCTGTTCTCCACGACTTCGCCATCACCAAAAACACTACGCGATTTTCCGAGATC	1044
QY	361	GlnValValPheLysLeuProGluMetIleArgGlyGlySerProValValTyrAspLys	380
DB	1045	CAAAATCCGGATGAGCCCAATCGAGATGCTGGTGGGGATCCCCGGTTAGCCGCGACGCT	1104
QY	381	AsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSerSerAsnIleLys	400
DB	1105	GGAAAGGTGCTCGCTCGGTCGGTGTGATCCTCGGTACGCGAAGACGAGTCCGAGATGAAG	1164
QY	401	TrpIleAspAlaProAspCysPheHisLeuTrpAsnAlaTrpGluGluProGlu	420
DB	1165	TGGTTGAGGTTCGGGTTTAAATGTGATACATTCATCAATGCGATGCGAGGAGGATGCG	1224
QY	421	ThrAspGluValValValIleGlySerCysMetThrProProAspSerIlePheAsnGlu	440
DB	1225	GGAGATACGGTGGTGTGCTGGTCCGAACATATATTCGTTGAACATACATCGTTGGAAGA	1284
QY	441	SerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsnLeuLysThrGlyGlu	460
DB	1285	ATGGAT---TTGATTCATGATCGGTCGAGAAAGTGAAGATTAATCTGAAACACAGAAATG	1341
QY	461	SerThrArgArgProIleIleSerAsnGluAspGlnGlnValAsnLeuGluAlaGlyMet	480
DB	1342	GTATCCCGGACCCGCTTCAACCCGG-----AATCTTGACTTTTCGGGTT	1386
QY	481	ValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeuAlaLeuAlaGluPro	500
DB	1387	TTAAATCCGGCTTTGTTGCCGTTAAAAACAGGTATATGTATGTGGAGTTGGTGTATCGG	1446
QY	501	TrpProLysValSerGlyPheAlaLysValAspLeuThrThrGlyGlu-----	516
DB	1447	ATGCCAAAGATCTCTGGTGTGCTCAAGCTAGATGTGTCTATCCGAAGACGACCGCTCGC	1506
QY	517	-----ValLysLysHisLeuTyrGlyAsnArgTyrGlyGlyGluProLeuPhe	533
DB	1507	GAATGCATAGTGTCTAGCCGATGTTGGGCTGTGTTGTACGGTGGTGAACCATCTTT	1566
QY	534	LeuProGlyGlu-----GlyGlyGluAspGluGlyTyrIleLeuCysPhe	549
DB	1567	GTACGTAGGAGCCAGCAACCCCGATGCGGATGAGGATGATGGGTATGTGTTCTTAT	1626
QY	550	ValHisAspGluLysThrTrpLysSerGluLeuGlnIleValAsnAlaValSer----	567
DB	1627	GTGCATACGAGAACACCGGTGAGTCAAGATTCGTGGTGTGAGCGCTAAGTACCGAGG	1686
QY	568	LeuGluValGluAlaThrValLysLeuProSerArgValProTyrGlyPheHisGlyThr	587



```
Db      1687 CTTGAGTCGTCGATCGCGTGAAGTCGCCACCGGGTACCATATGTTTTCATGGGCTC 1746
Qy      588 PheIleGlyAlaAspAepLeuAlaIys 596
Db      1747 TTTGTTAGAGAAAGTGACGTTTAACAAG 1773

RESULT 11
AAC36083
ID      AAC36083 standard; DNA; 1777 BP.
XX
AC      AAC36083;
XX
DT      17-OCT-2000 (first entry)
XX
DE      Arabidopsis thaliana DNA fragment SEQ ID NO: 12487.
XX
KW      Hybridisation assay; genetic mapping; gene expression control;
KW      protein identification; signal transduction pathway;
KW      metabolic pathway; promoter; termination sequence; ss.
XX
OS      Arabidopsis thaliana.
XX
EP1033405-A2.
XX
PD      06-SEP-2000.
XX
PF      25-FEB-2000; 2000EP-0301439.
XX
PR      25-FEB-1999; 99US-0121825.
PR      05-MAR-1999; 99US-0123180.
PR      09-MAR-1999; 99US-0123548.
PR      23-MAR-1999; 99US-0125788.
PR      25-MAR-1999; 99US-0126264.
PR      29-MAR-1999; 99US-0126785.
PR      01-APR-1999; 99US-0127462.
PR      06-APR-1999; 99US-0128234.
PR      08-APR-1999; 99US-0128714.
PR      16-APR-1999; 99US-0129845.
PR      19-APR-1999; 99US-0130077.
PR      21-APR-1999; 99US-0130449.
PR      23-APR-1999; 99US-0130510.
PR      28-APR-1999; 99US-0130891.
PR      30-APR-1999; 99US-0131449.
PR      30-APR-1999; 99US-0132048.
PR      30-APR-1999; 99US-0132407.
PR      04-MAY-1999; 99US-0132484.
PR      05-MAY-1999; 99US-0132485.
PR      06-MAY-1999; 99US-0132486.
PR      06-MAY-1999; 99US-0132487.
PR      07-MAY-1999; 99US-0132863.
PR      11-MAY-1999; 99US-0134256.
PR      14-MAY-1999; 99US-0134218.
PR      14-MAY-1999; 99US-0134219.
PR      14-MAY-1999; 99US-0134221.
PR      14-MAY-1999; 99US-0134221.
PR      14-MAY-1999; 99US-0134370.
PR      18-MAY-1999; 99US-0134760.
PR      19-MAY-1999; 99US-0134941.
PR      20-MAY-1999; 99US-0135124.
PR      21-MAY-1999; 99US-0135353.
PR      24-MAY-1999; 99US-0135629.
PR      25-MAY-1999; 99US-0136021.
PR      27-MAY-1999; 99US-0136392.
PR      28-MAY-1999; 99US-0136782.
PR      01-JUN-1999; 99US-0137222.
PR      03-JUN-1999; 99US-0137528.
PR      04-JUN-1999; 99US-0137502.
PR      07-JUN-1999; 99US-0137724.
PR      08-JUN-1999; 99US-0138094.
PR      10-JUN-1999; 99US-0138540.
PR      10-JUN-1999; 99US-0138847.
PR      14-JUN-1999; 99US-0139119.
PR      16-JUN-1999; 99US-0139453.
PR      16-JUN-1999; 99US-0139453.
PR      17-JUN-1999; 99US-0139453.
PR      18-JUN-1999; 99US-0139454.
PR      18-JUN-1999; 99US-0139455.
PR      18-JUN-1999; 99US-0139456.
PR      18-JUN-1999; 99US-0139457.
PR      18-JUN-1999; 99US-0139458.
PR      18-JUN-1999; 99US-0139459.
PR      18-JUN-1999; 99US-0139460.
PR      18-JUN-1999; 99US-0139461.
PR      18-JUN-1999; 99US-0139462.
PR      18-JUN-1999; 99US-0139463.
PR      18-JUN-1999; 99US-0139750.
PR      18-JUN-1999; 99US-0139763.
PR      21-JUN-1999; 99US-0139817.
PR      22-JUN-1999; 99US-0139899.
PR      23-JUN-1999; 99US-0140353.
PR      23-JUN-1999; 99US-0140354.
PR      24-JUN-1999; 99US-0140695.
PR      28-JUN-1999; 99US-0140823.
PR      29-JUN-1999; 99US-0140931.
PR      30-JUN-1999; 99US-0141287.
PR      01-JUL-1999; 99US-0141842.
PR      02-JUL-1999; 99US-0142154.
PR      02-JUL-1999; 99US-0142055.
PR      06-JUL-1999; 99US-0142390.
PR      08-JUL-1999; 99US-0142803.
PR      09-JUL-1999; 99US-0142920.
PR      12-JUL-1999; 99US-0142977.
PR      13-JUL-1999; 99US-0143542.
PR      14-JUL-1999; 99US-0143624.
PR      15-JUL-1999; 99US-0144005.
PR      16-JUL-1999; 99US-0144085.
PR      16-JUL-1999; 99US-0144086.
PR      19-JUL-1999; 99US-0144325.
PR      19-JUL-1999; 99US-0144331.
PR      19-JUL-1999; 99US-0144332.
PR      19-JUL-1999; 99US-0144333.
PR      19-JUL-1999; 99US-0144334.
PR      19-JUL-1999; 99US-0144335.
PR      20-JUL-1999; 99US-0144352.
PR      20-JUL-1999; 99US-0144632.
PR      20-JUL-1999; 99US-0144884.
PR      21-JUL-1999; 99US-0144814.
PR      21-JUL-1999; 99US-0145086.
PR      21-JUL-1999; 99US-0145088.
PR      22-JUL-1999; 99US-0145085.
PR      22-JUL-1999; 99US-0145087.
PR      22-JUL-1999; 99US-0145089.
PR      22-JUL-1999; 99US-0145192.
PR      23-JUL-1999; 99US-0145145.
PR      23-JUL-1999; 99US-0145218.
PR      23-JUL-1999; 99US-0145224.
PR      26-JUL-1999; 99US-0145276.
PR      27-JUL-1999; 99US-0145913.
PR      27-JUL-1999; 99US-0145918.
PR      27-JUL-1999; 99US-0145918.
PR      28-JUL-1999; 99US-0145951.
PR      02-AUG-1999; 99US-0146386.
PR      02-AUG-1999; 99US-0146388.
PR      02-AUG-1999; 99US-0146389.
PR      03-AUG-1999; 99US-0147038.
PR      04-AUG-1999; 99US-0147204.
PR      04-AUG-1999; 99US-0147302.
PR      05-AUG-1999; 99US-0147192.
PR      05-AUG-1999; 99US-0147260.
PR      06-AUG-1999; 99US-0147303.
PR      06-AUG-1999; 99US-0147416.
PR      09-AUG-1999; 99US-0147493.
PR      09-AUG-1999; 99US-0147935.
PR      10-AUG-1999; 99US-0148171.
PR      11-AUG-1999; 99US-0148319.
PR      12-AUG-1999; 99US-0148341.
```

```
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

## Alignment Scores:

```
Pred. No.: 1,048-81
Score: 939.00
Percent Similarity: 55.9%
Best Local Similarity: 37.03%
Query Match: 29.81%
DB: 21
```

```
Length: 1777
Matches: 207
Conservative: 106
Mismatch: 198
Indels: 48
Gaps: 13
```

US-09-758-269-6 (1-599) x AAC36083 (1-1777)

```
QY 64 LysGlnSerSerAsnSerProAlaIleValValLysProLys---AlaLysGluSerAsn 82
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 14 AAACCTCAGTGATGGCAGCATCATCATCTCAGTCCATCTCTAGACCCTCCAGGGTTTCTCC 73
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 83 ThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaAlaLeuAspAlaAlaGluGly 102
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 74 TCGAAGCTTCTCGATCTTCTCGAGAGACTTGTGTCAAGCTCATG-----118
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 103 PheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSerValGln 122
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 119 -----CACGATGCTTCTCTCCCTCTCCAC-----TAC 145
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 123 IleAlaGlyAsnPheAlaProVal---AsnGluGlnProValArgArgAsnLeuProVal 141
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 146 CTCTCAGGCAACTTCGCTCCCATCGGTGATGAACCTCTCTCCGCTCAAGGATCTCCCGTC 205
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 142 ValGlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyAlaAlaProLeu 161
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 206 CATGGATTTCCTCCGAATGCTTGAATGTTGAGGGTGGTCCAAACCCCAAG 265
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 162 HisGluProValThrGlyHisPhePheAspGlyAspGlyMetValHisAlaValLys 181
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 266 TTTGATGCTGCTGCTGGATATCAGTGGTTTGATGAGATGGATGATTCATGGGTACGC 325
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 182 PheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheValGln 201
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 326 ATCAAAAGATGGAAAGCTACTTATGTTCTCGATATGTTAAGACATCAGCTCTTAAGCAG 385
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 202 GluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGlyHisThr 221
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 386 GAAAAGTTCTTCGGAGCTGCCAAATTCATGAAG---ATTGGTGACCTTAAGGGGTTTTC 442
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 222 GlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaAlaGlyIleValAspProAla 241
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 443 GGATTGCTAATGGTCAATATCAACAGCTGAGAACGAGCTCAAATATTGACACACT 502
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 242 HisGlyThrGlyValAlaAlaAlaGlyLeuValTyrPheAsnGlyArgLeuLeuAlaMet 261
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 503 TATGAAATGGAACCTGCCAATACAGCACTCGTATATCACCATGGAACCTCTAGCATTA 562
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 262 SerGluAspAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysThrVal 281
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 563 CAGGAGGAGATAGCGTACGTACAAAGTTTGGAAAGATGGAGACCTGCAAACTCTT 622
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 282 GlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysValAsp 301
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 623 GGTATAATAGATTATGACAAGAGATTGACCCACTCTTCACTGCTCACCACCAAGTTGAC 682
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 302 ProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValSerLysProTyrLeuLys 321
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 683 CCGGTTACGGGTGAAATGTTTACATTCGGCTATTTCG---CATACGCCACCTTATCTCACA 739
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 322 TyrPheArgPheSerProAspGlyThrLysSerProAspValGluIleGlnLeuAspGln 341
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 740 TACAGAGTTATCTCGAAGATGGCATATGATGATGATGATGATGATGATGATGATGATGAT 799
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 342 ProThrMetMetHisAspPheAlaIleThrGluAsnPheValValValProAspGlnGln 361
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 800 CCTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 859
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 362 ValValPheLysLeuProGluMetIleArgGlyGlySerProValVal---TyrAspLys 380
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 860 ATGCACCTTCAGGCCCAAGGAAATGTTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 919
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 381 AsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSerSerAsnLys 400
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 920 ACAAAGAGCTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 979
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 401 TrpIleAspAlaProAspCysPheHisLeuTyrAsnAlaTrpGluProGlu 420
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 980 TGGTTTGAGCTTCCCACTGCTTTATTTTCCACACGCCCAATGCTTGGAA-----GAA 1033
```

Qy	421	ThrAspGluValValValIleGlySerCysMetThrProProAsp-----	435
Db	1034	GAGGATGAAGTCGTCCTCATCACTTGTGCTTGAGAAATCCAGATTCGATCGTCAGT	1093
Qy	436	SerIlePheAsnGluSerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsn	455
Db	1094	GGGAAAGTGAAGAAANAACCTCGAANAATTTGGCCACGAACTGTACGAAATGAGATTCCAC	1153
Qy	456	LeuLysThrGlyGluSerThrArgArgProIleIleSerAsnGluAspGlnValAsn	475
Db	1154	ATGAAACAGCGCTCAGCTTCTCAAAAACAACTATCCGCATCTCGG-----	1198
Qy	476	LeuGluAlaGlyMetValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyLeu	495
Db	1199	GTTGATTCCCCAGAATCAATGAGTGTACACCGGAAAGAAACAGAGATACGTATATGGA	1258
Qy	496	AlaLeuAlaGluProTyrProLysValSerGlyPheAlaLysValAsp-----	511
Db	1259	ACAATTCTGCACAGATATCGCAAGAGTTACCGGAATCATCAAGTTTGATTCGATGCGAGAA	1318
Qy	512	-----LeuThrThrGlyGluValLysLysHisLeuTyr-----	522
Db	1319	GCTGACACAGGAAAAACAAGTCTGGAGTAGGAGGTAAATCAAGGAATATATGACCTG	1378
Qy	523	GlyAspAsnArgTyrGlyGlyGluProLeuPheLeuProGlyGluGlyGlyGluGluAsp	542
Db	1379	GGAGAGGCAGATATGGTTCAGAGGCTATCTTATGTTCCGCGTGTAGACAGACAGAGAAGAC	1438
Qy	543	GluGlyTyrIleLeuCysPheValHisAspGluLysThrIleLysSerGluLeuGlnIle	562
Db	1439	GACGGTTACTTGATATTCTTTGTTTCATGATGAACAACACAGGGAATCATCGTGACGTGTG	1498
Qy	563	ValAsnAlaValSerLeuGluValGlu-----AlaThrValLysLeuProSerArgVal	580
Db	1499	ATAGACGCAAAAACAATGTCGGCTGAACCGGTGCAGTGTGGAGCTGCCCCACAGGGTC	1558
Qy	581	ProTyrGlyPheHisGlyThrPheIleGlyAlaAspAspLeuAlaLysGlnValVal	599
Db	1559	CCATATGGCTTCATGCTTGTTGTTGTTACAGAGAAACAATCCACGGAACAACTCTT	1615

## RESULT 12

RESOL. I.  
AAD09397

AA09397  
ID AAD09397 standard: cDNA: 1617 bp.

XX  
AC  
AAND00397.

AC  
XX  
DT

XX  
DE Arabidopsis thaliana neoxanthin cleavage enzyme. AtNCED4 cDNA.

Neoxanthin cleavage enzyme; ANCED4; abscisic acid; ABA; herbicide; stress tolerance; transgenic plant; plant breeding; antisense-therapy; plant growth protectant; ss.

XX Arabidopsis thaliana  
QS

XX	Key	Location/Qualifiers
FH	CDS	1..1617
FT		/*tag= a
FT		/product= "Arabidop

XX  
FBI 116704 AC

PN  
EPI16794-A2  
YY

XX  
PD  
18-JUL-2001

FD  
XX  
18-JUL-2001.

11 -JAN-2001: 2001EP-0300218-

XX  
FF 1007-1001, 2001

PR 13-JAN-2000: 2000

PR 11-JAN-2001; 2001

XX

DR WPI: 2001-400081/43.

DR P-PSDB; AAE04785.

XX

PT A DNA encoding a protein with a neoxanthin cleavage activity for  
PT producing transgenic plants with improved or decreased stress tolerance  
PT  
PT

yy  
Iy

XX  
PS Example 10: Page 39-43: 101pp: English:

XX 25

The invention relates to neoxanthin cleavage enzymes and their corresponding cDNA molecules. Neoxanthin cleavage enzyme plays a key role in endogenous abscisic acid (ABA) biosynthesis under drought stress. Neoxanthin cleavage enzyme is used for improving stress tolerance in a plant when expressed in a plant cell. The invention also relates to methods for increasing or decreasing stress tolerance in a plant by introducing the DNA into the plant, and a transgenic plant into which a neoxanthin cleavage enzyme is introduced. The improvement of stress tolerance in plants is useful, for example in plant breeding. Neoxanthin cleavage enzyme genes are useful for producing transgenic plants. An arid land can be improved by growing transformant weed for several years and then removing the weed by specifically lowering stress tolerance in the weed by inducing an inducible promoter. The present cDNA sequence encodes Arabidopsis thaliana neoxanthin cleavage enzyme, AtNCE4 protein.

U U

CC in the library using a cDNA of the CRD65 (Cowpea Responsive  
CC gene isolated from cowpea plant as a probe.  
CC  
XX Sequence 1617 BP: 480 A; 347 C; 369 G; 421 T; 0 others;  
SO

Program	Length	Score
1.13e-81	1617	

55  
56

Percent Similarity:	55.81%	Conservative:	105
Best Local Similarity:	37.03%	Mismatches:	199
Query Match:	29.78%	Indels:	48
DR:	22	Gaps:	13

```
Db 439 GGAATGCTCTAATGCTCAATGCTCAACAGCTGAGAACGAAAGCTCAAAATATTGGACAACT 498
Qy 242 HsGlyThrGlyValAlaAsnAlaGlyLeuValTyrPheAsnGlyArgLeuLeuAlaMet 261
Db 499 TATGAAATGGAACCTGCCAATACAGCACTCGTATATACCATGGAACACTTCAGCAATTA 558
Qy 262 SerGluAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysThrVal 281
Db 559 CAGGAGGCGCATGAAGCGGTACGTGCATCAAAAGTTTGGAAAGATGAGACCTGCAAACTCTT 618
Qy 282 GlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysValAsp 301
Db 619 GGTATAATAGATTATGACAAGAGATTGACCCACTCTCTTCACTGCTCACCACAAAGTTGAC 678
Qy 302 ProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValSerLysProTyrLeuLys 321
Db 679 CCGGTTACGGGTGAATGTTTACATTCGGCTATTGCG---CATACGCCACTTATCTCACA 735
Qy 322 TyrPheArgPheSerProAspGlyThrLysSerProAspValGluIleGlnLeuAspGln 341
Db 736 TACAGAGTTATCTCGAAAGATGCGCATTTATGCAATGACCCAGTCCCAATTAATATCAGAG 795
Qy 342 ProThrMetMetHisAspPheAlaIleThrGluAsnPheValValProAspGlnGln 361
Db 796 CTTATCATGATCATGATTTTCTATTCTAGAGATTATGCAATCTTCAATGATCTTCTCT 855
Qy 362 ValValPheLeuProGluMetIleArgGlyGlySerProValVal---TyrAspLys 380
Db 856 ATGCACCTTCAGGCCAAGCAATGTTGAAGAGAGAGAAATGATATACTACTTCATTCATCCC 915
Qy 381 AsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSerSerAsnLys 400
Db 916 ACAAAAAAGCGCTCGTTTGGTCTCTTCCACGCTATGCCAAGATGCAACTTATGATTAGA 975
Qy 401 TrpIleAspAlaProAspCysPheCysPheHisLeuTrpAsnAlaTrpGluGluProGlu 420
Db 976 TGGTTTGAGCTTCCCAACGCTTTATTTTCCCAACGCGCAATGCTTGGGAA-----GAA 1029
Qy 421 ThrAspGluValValIleGlySerCysMetThrProProAsp----- 435
Db 1030 GAGGATGAAGTCTCTCATCACTTGTCTCTTGAGATCGAGATCTTGACATGTCAGT 1089
Qy 436 SerIlePheAsnGluSerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsn 455
Db 1090 GGGAAAGTGAAGAAAAAAGCTCGAAATTTTGGCAACGAACTGTACGAAATGAGATTCAAC 1149
Qy 456 LeuLysThrGlyGluSerThrArgArgProIleIleSerAsnGluAspGlnValAsn 475
Db 1150 ATGAAAAAGGGCTCAGCTTCTCAAAAAAACTATCCGCACTCGCG----- 1194
Qy 476 LeuGluAlaGlyMetValAsnArgAsnMetLeuGlyArgGlyThrLysPheAlaTyrLeu 495
Db 1195 GTTGATTTCCTCCAGATCATGAGTGTACACCGGAAGAAACAGAGATACGTATATGA 1254
Qy 496 AlaLeuAlaGluProTrpProLysValSerGlyPheAlaLysValAsp----- 511
Db 1255 ACAATTCTGGACAGTATCGAAAGGTTACCGGAATCATCAAGTTTGATCGCATCGAGAA 1314
Qy 512 -----LeuThrThrGlyGluValLysLysHisLeuTyr----- 522
Db 1315 CTTGAGACAGGGAAGAAAGTGTGGAAGTAGAGGTAATATCAAGGAATATATACACCTG 1374
Qy 523 GlyAspAsnArgTyrGlyGlyGluProLeuPheLeuProGlyGlyGlyGluGluAsp 542
Db 1375 CGAAGAGCGAGATATGGTTCAGAGGCTATCTATGTTCCGCTGAGACAGCAGAAAGAC 1434
Qy 543 GluGlyTyrIleLeuCysPheValHisAspGluLysThrTrpLysSerGluLeuGlnIle 562
Db 1435 GACGGTTACTTGTATATCTTTCTTCTCATGATGAAACACAGGGAATCATCGTGTGCTGTG 1494
Qy 563 ValAsnAlaValSerLeuValGlu-----AlaThrValLysLeuProSerArgVal 580
```

```
Db 1495 ATAGACGCCAAAAACAACTATGCGCTGAAACCGGTGGCAGTGGTGGAGCTGCCGACAGGTC 1554
Qy 581 ProTyrGlyPheHisGlyThrPheIleGlyAlaAspAspLeuAlaLysGlnValVal 599
Db 1555 CCATATGGCTTCCATCGCTTGTGTTGTACAGAGGAACAACTCCAGGAACAAACTCTT 1611

RESULT 13
AAD09402
ID AAD09402 standard; cDNA; 1617 BP.
XX
AC AAD09402;
XX
DT 10-SEP-2001 (first entry)
XX
DE Arabidopsis thaliana neoxanthin cleavage enzyme cDNA.
XX
KW Neoxanthin cleavage enzyme; abscisic acid; ABA; herbicide;
KW stress tolerance; transgenic plant; plant breeding; antisense-therapy;
KW plant growth protectant; ss.
XX
OS Arabidopsis thaliana.
XX
FH Key
CDS Location/Qualifiers
FT 1..1617
FT /*tag= a
FT /product= "Arabidopsis neoxanthin cleavage enzyme"
XX
PN EP1116794-A2.
PD 18-JUL-2001.
XX
PF 11-JAN-2001; 2001EP-0300218.
XX
PR 13-JAN-2000; 2000JP-0010056.
PR 11-JAN-2001; 2001JP-0003476.
XX
PA (RIKE ) RIKEN KK.
PI
PI Iuchi S, Kobayashi M, Shinozaki K;
XX
DR WPI: 2001-400081/43.
DR P-PSDB; AAE04790.
XX
PT A DNA encoding a protein with a neoxanthin cleavage activity for
PT producing transgenic plants with improved or decreased stress tolerance
PS
PS Disclosure; Page 74-77; 101pp; English.
XX
CC The invention relates to neoxanthin cleavage enzymes and their
CC corresponding cDNA molecules. Neoxanthin cleavage enzyme plays a key
CC role in endogenous abscisic acid (ABA) biosynthesis under drought stress.
CC Neoxanthin cleavage enzyme is used for improving stress tolerance in a
CC plant when expressed in a plant cell. The invention also relates to
CC methods for increasing or decreasing stress tolerance in a plant by
CC introducing the DNA into the plant, and a transgenic plant into which a
CC neoxanthin cleavage enzyme is introduced. The improvement of stress
CC tolerance in plants is useful, for example in plant breeding. Neoxanthin
CC cleavage enzyme genes are useful for producing transgenic plants. An arid
CC land can be improved by growing transformant weed for several years and
CC then removing the weed by specifically lowering stress tolerance in the
CC weed by inducing an inducible promoter. The present cDNA sequence encodes
CC Arabidopsis thaliana neoxanthin cleavage enzyme related to the invention.
XX
SQ Sequence 1617 BP; 480 A; 348 C; 368 G; 421 T; 0 other;
```

```
Alignment Scores:
Pred. No.: 1,426-81 Length: 1617
Score: 937.00 Matches: 207
Percent Similarity: 55.81% Conservative: 105
Best Local Similarity: 37.03% Mismatches: 199
Query Match: 29.75% Indels: 48
DB: 22 Gaps: 13
```

```
US-09-758-269-6 (1-599) x AAD09402 (1-1617)
Qy 64 LysGlnSerSerAsnSerProAlaIleValValIysProLys---AlaLysGluSerAsn 82
Dy 10 AAACCTCAGTGATGGCAGCATCATCTCATCTCCATCCATAGACCTCCAGGGTTCTCC 69
Qy 83 ThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaLeuAspAlaAlaGluGly 102
Dy 70 TCGAAGCTTCCTCGATCTCTCGAGAGACTTCTCGTCAAGCTCATG----- 114
Qy 103 PheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSerValGln 122
Dy 115 -----CAGGATCTCTCTCCCTCTCCAC-----TAC 141
Qy 123 IleAlaGlyAsnPheAlaProVal---AsnGluGlnProValArgAsnLeuProVal 141
Dy 142 CTCTCAGGCAACTTCGCTCCCATCGTGATGAAACTCTCCCGTCAAGGATCTCCCGTC 201
Qy 142 ValGlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyAlaAsnProLeu 161
Dy 202 CATGGATTTCTCCGAAATGCTTGAATGGTGAATTTGTAGGGTTGTGTCGAACCCCAAG 261
Qy 162 HisGluProValThrGlyHisPhePheAspGlyAspGlyMetValHisAlaValLys 181
Dy 262 TTGTGATGCTGCTGGATATCATCTGTTGATGGAGATGGATGATTCATGGGTAGCG 321
Qy 182 PheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheValGln 201
Dy 322 ATCAAGATGGAAAGCTACTTATGTTCTCGATATGTTAAGACATCATCAGCTCTTAAGCAG 381
Qy 202 GluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGlyHisThr 221
Dy 382 GAAGAGTTCTTCGAGCTGCCAAATTCATGAAG---ATTGTTGACCTTAAAGGGTTTTC 438
Qy 222 GlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaAlaGlyIleValAsnProAla 241
Dy 439 GGATTTGTAATGGTCAATATCAACAGCTGAGAGAGAGCTCAAAATATTGGACAACT 498
Qy 242 HisGlyThrGlyValAlaAsnAlaGlyLeuValTyrPheAsnGlyArgLeuLeuAlaMet 261
Dy 499 TATGNAATGAACTGCCAATACAGCACTCGTATATACCATGGAAGAACTTCTAGCATTA 558
Qy 262 SerGluAspAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysThrVal 281
Dy 559 CAGGAGGAGATAGACCGCTACGTATCAAAAGTTTGGAGAGATGGAGACCTGCCAACTCT 618
Qy 282 GlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysValAsp 301
Dy 619 GGTATATAGATTATGACAAGAGATTGACCACTCTCTCTCTCTCTCTCTCTCTCTCTCT 678
Qy 302 ProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValSerLysProTyrLeuLys 321
Dy 679 CCGGTTACGGGTGAAATGTTTACATTCGGCTATTGCG---CATAGCCCACTTATCTCACA 735
Qy 322 TyrPheArgPheSerProAspGlyThrLysSerProAspValGluIleGlnLeuAspGln 341
Dy 736 TACAGAGTTATCTCGAAAGATGGCATTATGATGACCCAGTCCCAATTTACTATATACAG 795
Qy 342 ProThrMetMetHisAspPheAlaIleThrGluAsnPheValValProAspGlnGln 361
Dy 796 CCTATCATGATGATGATTTGTTATTTACTGAGACTTATGCAATCTTCATGATCTTCT 855
Qy 362 ValValPheLysLeuProGluMetIleArgGlyGlySerProValVal---TyrAspLys 380
Dy 856 ATGCACCTTCAGGCCAAAGAAATGTTGAAGAGAGAAATGATATATCTCATTTGATCCC 915
Qy 381 AsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSerAsnIleLys 400
Dy 916 ACAAAAAAGGCTCGTTGGTCTTCCTCCGCGCTATGCCAAGGATGAACCTTATGATTAGA 975
Qy 401 TrpIleAspAlaProAspCysPheCysPheHisLeuTrpAsnAlaTrpGluGluProGlu 420
```

```
Dy 976 TGGTTTCAGCTTCCCAACTGCTTTATTTTCCAAACCAATGCTTGGGAA-----GAA 1029
Qy 421 ThrAspGluValValIleGlySerCysMetThrProProAsp----- 435
Dy 1030 GAGGATGAAGTCGCTCATCTGCTTGTGAGAAATCCAGATCTTGGATGGTCAGT 1089
Qy 436 SerIlePheAsnGluSerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsn 455
Dy 1090 GGGAAAGTGAAGAAAAAACTCGAAAAATTTTGGCAACGAACCTGTACGAATGAGATTCAAC 1149
Qy 456 LeuLysThrGlyGluSerThrArgProIleIleSerAsnGluAspGlnGlnValAsn 475
Dy 1150 ATGAAAAAGGCTCAGCTTCTCAAAAAAACTATCCGCATCTGCG----- 1194
Qy 476 LeuGluAlaGlyMetValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeu 495
Dy 1195 GTTGATTTCCCGCAATCAATGATGCTACACCGGAAAGAAACAGAGATACGTATATGGA 1254
Qy 496 AlaLeuAlaGluProTrpProLysValSerGlyPheAlaLysValAsp----- 511
Dy 1255 ACAATTCGGACAGTATCGCAAGGTTACCGGAATCATCAAGTTTGTCTGCATGCAGAA 1314
Qy 512 -----LeuThrThrGlyGluValLysLysHisLeuTyr----- 522
Dy 1315 GCTGAGACAGGAAAAAGAAATGCTGGAAGTAGGAGGTAAATATCAAGGAATATATGACCTG 1374
Qy 523 GlyAspAsnArgTyrGlyGluProLeuPheLeuProGlyGluGlyGlyGluGluAsp 542
Dy 1375 GGAGAAAGCAGATATGTTTCCAGAGGCTATCTATGTTCCGCTGAGACGACGAAAGAGAC 1434
Qy 543 GluGlyTyrIleLeuCysPheValHisAspGluLysThrTrpLysSerGluLeuGlnIle 562
Dy 1435 GACGGTTACTGTATGATTTCTTGTTCATGATGAACACAGGGAATCATGCGTACTGTG 1494
Qy 563 ValAsnAlaValSerLeuGluValGlu-----AlaThrValLysLeuProSerArgVal 580
Dy 1495 ATAGACGCAAAACAATGTCGGCTGAACCGGTGCGAGTGGTGGAGCTGCCGACAGGGTC 1554
Qy 581 ProTyrGlyPheHisGlyThrPheIleGlyAlaAspAspLeuAlaLysGlnValVal 599
Dy 1555 CCATATGGCTTCCATGCTGTTTGTGTACAGAGGAACAACCTCCAGGAACAACCTCTT 1611
RESULT 14
AAC56678
ID AAC56678 standard; DNA; 443 BP.
XX AAC56678;
AC AAC56678;
XX 25-JAN-2001 (first entry)
DE Eucalyptus grandis transcription factor DNA sequence #549.
XX Plant; transcription factor; gene expression; eucalyptus; pine; acacia;
XX poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;
XX basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;
XX homeodomain zipper; LIM domain; AP2; EREBs; zinc finger domain;
XX type 2 Cys2His2; CCAAT box element; MYB; ss.
OS Eucalyptus grandis.
XX WO200053724-A2.
XX 14-SEP-2000.
XX 09-MAR-2000; 2000WO-US061112.
XX 11-MAR-1999; 99US-0266513.
XX 18-AUG-1999; 99US-0149485.
XX (GENE-) GENESIS RES & DEV CORP LTD.
XX (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX Wood M, McGrath A, Shenk MA, Glenn M;
PI
```

XX WPI; 2000-579369/54.  
XX New isolated polynucleotide encoding a plant transcription factor for  
PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,  
PT having modified gene expression or modified activity of a polypeptide  
PT  
XX  
XX Claim 1; Page 490; 747pp; English.  
XX The present invention relates to novel plant transcription factors from  
CC Eucalyptus grandis or Pinus radiata. The present sequence is the coding  
CC sequence for one such transcription factor. The transcription factor may  
CC be used to produce a plant having modified gene expression such as a  
CC woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or  
CC mahogany species or to modify the activity of a polypeptide in a plant.  
CC The transcription factors of the present invention are members from the  
CC following families of regulatory proteins: bZIP, bZIP family of G-box  
CC binding factors, basic helix-loop-helix zipper,  
CC homeotic/homeodomain/homeobox/WADS, homeodomain zipper, LIM domain, AP2  
CC and EREBs, zinc finger domains of type 2 Cys2His2, CCAAT box elements  
CC and MYB.  
XX  
SQ Sequence 443 BP; 71 A; 177 C; 117 G; 77 T; 1 other;  
Alignment Scores:  
Pred. No.: 3.38e-50 Length: 443  
Score: 608.00 Matches: 115  
Percent Similarity: 87.76% Conservative: 14  
Best Local Similarity: 78.23% Mismatches: 18  
Query Match: 19.30% Indels: 0  
DB: 21 Gaps: 0  
US-09-758-269-6 (1-599) x AAC56678 (1-443)  
QY 208 ProValPheProLysAlaIleGlyLeuLeuHisGlyHisThrGlyIleAlaArgLeuMet 227  
DB 3 CCGGCTTTCCNCAAGCCATCGCGAGCTCCACGGCCACTCGGATCGCGGCTCATG 62  
QY 228 LeuPheTyrAlaArgAlaAlaGlyIleValAspProAlaHisGlyThrGlyValAla 247  
DB 63 CTCTTCTACGCCCGGAGGCTCTTCGGCTCTGTCGACCCGGAATGGATGGCGTCGCG 122  
QY 248 AsnAlaGlyLeuValTyrPheAsnGlyArgLeuLeuAlaMetSerGluAspLeuPro 267  
DB 123 AACGGCGGCTCGTGACTTCGACGGCCACCTCTCTCGGATGTCGAGGACGACCTCCC 182  
QY 268 TyrGlnValGlnIleThrProAsnGlyAspLeuLysThrValGlyArgPheAsp 287  
DB 183 TACCACGTGCGCTCACGCGCTCCGCGACCTCGAGACCGTCCGCGCTACGACTTCGCC 242  
QY 288 GlyGlnLeuGluSerThrMetIleAlaHisProLysValAspProGluSerGlyGluLeu 307  
DB 243 GGCCAGCTCGACTCTCCGATGATGCCACCCGAAGATCGACCCGGCTTCGCGGAGATG 302  
QY 308 PheAlaLeuSerTyrAspValValSerLysProTyrLeuLysTyrPheArgPheSerPro 327  
DB 303 TTCGCCCTCAGCTACGACGTGTCGCGAAGCGGTACTCTCACTCAAGTACTTCGATTC 362  
QY 328 AspGlyThrLysSerProAspValGluLeuGlnLeuAspGlnProThrMetMetHisasp 347  
DB 363 GACGCGGAGAAGTCCCGGAGCGTGAGATCCCGCTGGCTGAGCCGACCATGATGCACGAT 422  
QY 348 PheAlaIleThrGluAsnPhe 354  
DB 423 TTCGCCATACCGAGCGGCTTT 443  
RESULT 15  
ID AAC82706/c  
XX AAC82706 standard; DNA; 492 BP.  
XX  
AC AAC82706;  
XX

DT 15-MAR-2001 (first entry)  
XX  
DE Rice abscisic acid synthesis associated DNA SEQ ID NO 5.  
XX  
KW Rice; abscisic acid synthesis; VP14-like gene; drought resistance; ds.  
XX  
OS Oryza sativa.  
XX  
PN WO200071727-A1.  
XX  
PD 30-NOV-2000.  
XX  
PF 25-MAY-1999; 99WO-JP02734.  
XX  
PR 25-MAY-1999; 99WO-JP02734.  
XX  
PA (NORO) JAPAN MIN AGRIC FORESTRY & FISHERIES.  
PA (BIOO-) BIO-ORIENTED TECHNOLOGY RES ADVANCEMENT.  
XX  
PI Hirochika H, Sakamoto K;  
XX  
DR WPI; 2001-032042/04.  
XX  
XX Oligonucleotide encoding gene for regulating abscisic acid synthesis in  
PT plants, useful for constructing e.g. genetically-modified rice with  
PT drought resistance and ear-germination resistance  
XX  
PS Disclosure; Page 50; 55pp; Japanese.  
XX  
CC This invention describes a novel polynucleotide sequence (I) which  
CC encodes a protein capable of regulating the synthesis of abscisic acid.  
CC The invention also describes (1) an oligonucleotide encoding a protein  
CC of a gene obtained by controlling the expression of a VP14-like gene;  
CC (2) a vector containing the oligonucleotide ligated operably to the  
CC regulation sequence; (3) a plant transformed with the vector; and  
CC (4) a method for regulating abscisic acid synthesis in a plant including  
CC the transfer of the above oligonucleotide to it. The gene is useful for  
CC constructing drought resistant rice.  
XX  
SQ Sequence 492 BP; 90 A; 126 C; 186 G; 88 T; 2 other;  
Alignment Scores:  
Pred. No.: 1.45e-48 Length: 492  
Score: 592.00 Matches: 115  
Percent Similarity: 90.28% Conservative: 15  
Best Local Similarity: 79.86% Mismatches: 13  
Query Match: 18.79% Indels: 2  
DB: 22 Gaps: 0  
US-09-758-269-6 (1-599) x AAC82706 (1-492)  
QY 210 PheProLysAlaIleGlyLeuLeuHisGlyHisThrGlyIleAlaArgLeuMetLeuPh 229  
DB 434 TTTCTTAAGCGATAGGTGAGCTTCATGGCCACTCCGGGCGATCGCGCCCTTGCTCTGT 375  
QY 229 eTyrAlaArgAlaAlaGlyIleValAspProAlaHisGlyThrGlyValAlaAsnAl 249  
DB 374 CTACGGCGCGCGGCTTC-GGCTCTCTCGACCGCTCACACGGCAGCGGCTGCGCAAGCG 316  
QY 249 aGlyLeuValTyrPheAsnGlyArgLeuLeuAlaMetSerGluAspLeuProTyrGl 269  
DB 315 CGGCTCATCTACTTCAACGGCAGGCTCTCGCCATGTGCGGAGGACGACCTCCCTACCA 256  
QY 269 nValGlnIleThrProAsnGlyAspLeuLysThrValGlyArgPheAspPheAspGly 289  
DB 255 GGTGCGGCTCACCGCGGACCTCGAGACCGTCCGCGGCTCGACCTTCGACGGGCA 196  
QY 289 nLeuGluSerThrMetIleAlaHisProLysValAspProGluSerGlyGluLeuPheAl 309  
DB 195 GCTGGGTGCGGCTATGTCGCGGACCCCAAGCTCGACCGGCGGAGAGCTCCAGCG 136  
QY 309 aLeuSerTyrAspValValSerLysProTyrLeuLysTyrPheArgPheSerProAspGl 329

Db	135	GCTCAGCTACGACGTGATCAAGAAGCCGTACCTCAAGTACTTCTTACTTCGGGCCCGACGG	76
Qy	329	YThrlYsSerProAspValGluIleGlnLeuAspGlnProThrMetMetHisAspPheAl	349
Db	75		
Qy	349	aileThrGlu	352
Db	15	AATTACTGAG	6

Search completed: November 15, 2003, 15:07:18  
Job time : 522 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 15, 2003, 12:55:19 ; Search time 103 Seconds  
(without alignments)  
2566.879 Million cell updates/sec

Title: US-09-758-269-6  
Perfect score: 3150  
Sequence: 1 MASFATATAVSGRWLGGNHT.....VPYGFHGTFIGADDLAKQV 599

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-O=/cn2\_1/uspto.spool/US09758269/runat.14112003.135757.28769/app.query.fasta.1.775  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LCOFCL=0  
-LOOPEXT=0 -UNITS=Bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09758269 @CN 1.1.56 @runat.14112003.135757.28769 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:  
1: /cn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cn2\_6/ptodata/2/ina/PTUS\_COMB.seq:\*  
6: /cn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	494.5	15.7	4403765	3	US-09-103-840A-2
2	494.5	15.7	4411529	3	US-09-103-840A-1
3	475.5	15.1	4403765	3	US-09-103-840A-2
4	475.5	15.1	4411529	3	US-09-103-840A-1
5	419	13.3	32679	4	US-08-976-063E-1
6	411.5	13.1	1518	4	US-08-976-063E-21
7	211	6.7	2629	1	US-08-200-807-1
8	211	6.7	2629	1	US-08-488-305A-1
9	210	6.7	1724	3	US-09-385-259-1
10	210	6.7	1724	4	US-09-645-370-1
11	186	5.9	36063	4	US-08-311-731A-140
12	127.5	4.0	4242	4	US-09-252-991A-7056

13	127.5	4.0	10023	4	US-09-252-991A-6997	Sequence 6997, Ap
14	123.5	3.9	5163	3	US-08-700-651-1	Sequence 1, Appli
15	123.5	3.9	5163	3	US-08-928-361B-4	Sequence 4, Appli
16	123.5	3.9	5163	3	US-09-588-995A-4	Sequence 4, Appli
17	123.5	3.9	5318	3	US-08-700-651-2	Sequence 2, Appli
18	123.5	3.9	5318	3	US-08-928-361B-3	Sequence 3, Appli
19	123.5	3.9	5318	4	US-09-588-995A-3	Sequence 3, Appli
20	116.5	3.7	5511	3	US-08-928-361B-2	Sequence 2, Appli
21	116.5	3.7	5511	4	US-09-588-995A-2	Sequence 2, Appli
22	116.5	3.7	7334	3	US-08-928-361B-1	Sequence 1, Appli
23	116.5	3.7	7334	4	US-09-588-995A-1	Sequence 1, Appli
24	115	3.7	1389	4	US-09-252-991A-11721	Sequence 11721, A
25	115	3.7	2322	4	US-09-252-991A-11519	Sequence 11519, A
26	115	3.7	47981	4	US-09-679-279-1	Sequence 1, Appli
27	110	3.5	1635	4	US-09-252-991A-2493	Sequence 2493, Ap
28	106	3.4	1239	3	US-09-461-697-76	Sequence 76, Appl
29	106	3.4	1825	3	US-09-461-697-75	Sequence 75, Appl
30	106	3.4	1953	4	US-09-252-991A-759	Sequence 759, App
31	105.5	3.3	1416	4	US-09-107-532A-2051	Sequence 2051, Ap
32	105.5	3.3	4508	5	PCT-US93-06251-34	Sequence 34, Appl
33	105	3.3	1172	3	US-08-861-774E-17	Sequence 17, Appl
34	104.5	3.3	1389	4	US-09-328-352-3002	Sequence 3002, Ap
35	104	3.3	1341	4	US-09-350-756-2	Sequence 2, Appli
36	103	3.3	1479	4	US-09-252-991A-4180	Sequence 4180, Ap
37	103	3.3	1575	4	US-09-252-991A-2009	Sequence 2009, Ap
38	103	3.3	1581	4	US-09-252-991A-654	Sequence 654, App
39	103	3.3	3102	4	US-09-252-991A-4429	Sequence 4429, Ap
40	102.5	3.3	2607	2	US-08-907-166-1	Sequence 1, Appli
41	102.5	3.3	2607	4	US-09-391-340-1	Sequence 1, Appli
42	102.5	3.3	3257	4	US-09-585-173B-39	Sequence 39, Appl
43	102.5	3.3	14672	4	US-08-961-521-111	Sequence 111, App
44	102	3.2	879	4	US-09-252-991A-12072	Sequence 12072, A
45	102	3.2	3396	4	US-09-252-991A-14676	Sequence 14676, A

ALIGNMENTS

RESULT 1  
US-09-103-840A-2  
; Sequence 2, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; TITLE OF INVENTION: TUBERCULOSIS  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; TYPE: DNA  
; LENGTH: 4403765  
; ORGANISM: Mycobacterium tuberculosis  
; FEATURE:  
; OTHER INFORMATION: CDC 1551  
; OTHER INFORMATION: "n" bases at various positions throughout the sequence  
; OTHER INFORMATION: represent a, t, c or g  
US-09-103-840A-2

Alignment Scores:  
Pred. No.: 4,228-39 Length: 4403765  
Score: 494.50 Matches: 180  
Percent Similarity: 40.18% Conservative: 94  
Best Local Similarity: 26.39% Mismatches: 251  
Query Match: 15.70% Indels: 157  
DB: 3 Gaps: 27

US-09-758-269-6 (1-599) x US-09-103-840A-2 (1-4403765)

```
Qy 1 MetAlaSerPheThrAlaThr---AlaAlaValSerGly-----Arg 13
Db 751637 ATGGCCAGAGTTCGGCAGCGGATCGCGCGTCTGCGGAGCGGTTCAGCGCGACATCGAAC 751696
Qy 14 TrpLeuGlyGlyAenHisThrGlnProProLeuSerSerSerGlnSerSerAspLeuSer 33
Db 751697 TGGCGTACCCCTTCTTCGGCTATGGCGCGATGATCGCGCATCCCGCGCAAAATGGGTG 751756
Qy 34 TyrCysSer-----SerLeuProMetAlaSerArgValThrArgLysLeuAsnVal--- 50
Db 751757 TACACCGCATCGTCGAGTGCTGCTGCGCGCCACCTTTCGGGGTTCGAGCGCGTCG 751816
Qy 51 -----SerSerAlaLeuHisThrProAlaLeuHisPheProLysGlnSerSer 67
Db 751817 GCCCGTGATCGTGGAGCAGTCGACGCCGCGGTGC-----AGCAGCTCGTCG 751864
Qy 68 AsnSerProAlaIleValVal-----LysProLysAlaLysGlnSerAsn 82
Db 751865 CGAACACCGGTCTCGAGGTCTATCTTCCTCCATGTTCTCACCAAGGGGTACCGTTCCTCAAT 751924
Qy 83 ThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaLeuAspAlaAlaGluGly 102
Db 751925 ATCAGTGAATAACAATTTATAGAGATCGGCATGACACCGCACAGCCGCGAA--- 751981
Qy 103 PheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSerValGln 122
Db 751982 -----TCCCAAAACCCCATATCTCGAG 752002
Qy 123 IleAlaGlyAenPheAlaProValAsnGlnProValArgArgAsnLeuProValVal 142
Db 752003 -----GGCTCTTCGGCGCGGTGAGCACCGAGGTAACTGCCACCGACCTTCGCGTCACC 752056
Qy 143 GlyLysLeuProAspSerIleGlyValTyrValArgAsnGlyAlaAsnProLeuHis 162
Db 752057 GCCCGCATTCGGGAACACCTACGACGCGGTATCTGCGTAAACGCGCCCAACCCGCGCG 752116
Qy 163 Glu-----ProValThrGlyHisPhePheAspGlyAspGlyMetValHisAlaVal 180
Db 752117 GAGGTGACCCCGGCCACC---TACACTGGTTTACCGCGCAGCGCATGGTGCACGGAGTC 752173
Qy 181 LysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheVal 200
Db 752174 CGCTGCGCGACGGGAAGGCC-----CGCTGGTAT 752203
Qy 201 GlnGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGlu----- 216
Db 752204 CGCAATCGTGGTCCGACACCGCGGTGTGCGCGCGCTTCGGCGAGCCCATTCGGCC 752263
Qy 217 ---LeuHisGlyHisThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAla 235
Db 752264 CGGCTCACCCCGCGACCGGGATTATCGAG----- 752293
Qy 236 GlyIleValAspProAlaHisGlyThrGlyValAlaAenAlaGlyLeuValTyrPheAsn 255
Db 752294 -----GGCGGTCCCAACACCAACGCTGCTGACCCACGCC 752326
Qy 256 GlyArgLeuLeuAlaMetSerGluAspAspLeuProTyrGlnValGlnIleThrProAsn 275
Db 752327 GGACGCACTTCGGCTCGGTGGAGCGCGCGT-----GTCAACTACGAACTCACC 752377
Qy 276 GlyAspLeuLysThrValGlyArgPheAspGlyGlnLeuGlnSerThrMetIle 295
Db 752378 GATGAGCTGGACACCGTGGACCTGTGACTTCACGCGCACCTTCGACCGCGGTTCACACC 752437
Qy 296 AlaHisProLysValAspProGluSerGlyGluLeuPheAlaLeuSerTyrAspValVal 315
Db 752438 GCCCATCCGACGTGATCCGACACCGGTGAAGTGCACGCGGTCTCTACTCTGCTCGCC 752497
Qy 316 SerLysProTyrLeuLysTyrPheArgPheSerProAspGlyThrLysSerProAspVal 335
Db 752498 CGCGACACAGAGTGCAGTACTCTGCTGATCGGCACCGACCGACCGTTCGTCGACCGGTT 752557
Qy 336 GluIleGlnLeuAspGlnProThrMetMetHisAspPheAlaIleThrGluAsnPheVal 355
```

## RESULT 2

US-09-103-840A-1  
; Sequence 1, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.

```

; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.: 4,238-39 Length: 4411529
Score: 494.50 Matches: 180
Percent Similarity: 40.18% Conservative: 94
Best Local Similarity: 26.39% Mismatches: 251
Query Match: 15.70% Indels: 157
DB: 27

US-09-758-269-6 (1-599) x US-09-103-840A-1 (1-4411529)
QY 1 MetAlaSerPheThrAlaThr---AlaAlaValSerGly-----Arg 13
DB 749677 ATGGCCAGCAGTCTGGCGACCGATCGGCGGTCGCGGACCGTCAGCGCGATCGAAC 749736.
QY 14 TrpLeuGlyGlyAsnHisThrGlnProProLeuSerSerSerGlnSerSerAspLeuSer 33
DB 749737 TGGCGTAGCCTTCTCGGCTATGGCGGATCAGTCCGCGCATCCCGCGAAATGGGTG 749796
QY 34 TyrCysSer-----SerLeuProMetAlaSerArgValThrArgLysLeuAsnVal--- 50
DB 749797 TACACCGCATCGTCAGGTGCTCTCGCGCGGCACCTTCGCGGTCTGCGAGCGGTG 749856
QY 51 -----SerSerAlaLeuHisThrProProAlaLeuHisPheProLysGlnSer 67
DB 749857 GCGCGTGTATCTCGAGCAGTCGACCGCGCGGTG-----AGCAGCTCGTCG 749904
QY 68 AsnSerProAlaIleValVal-----LysProLysAlaLysGluSerAsn 82
DB 749905 CGAACACCGTCTCGGAGTCTATCTTCCATGTTCTCACCAAGGCGGTACCGTTCCAAT 749964
QY 83 ThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaLeuAspAlaGluGly 102
DB 749965 ATCACTGAAATAACAATGTTATAGGAGATCGCATCCACCGCACCGCAAGCCCGAA--- 750021
QY 103 PheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSerValGln 122
DB 750022 -----TCCAAACCCATATCTCGAG 750042
QY 123 IleAlaGlyAsnPheAlaProValAsnGluGlnProValArgAsnLeuProValVal 142
DB 750043 -----GGCTTCTCGCGCGGTGAGCACCGAGTAACCTGCCACCGACCTCGCGTCACC 750096
QY 143 GlyLysLeuProAspSerIleLysGlyValThrValArgAsnGlyAlaAsnProLeuHis 162
DB 750097 GCGCGCATTCGCGAAACACCTCGACGGCGGTATCTCGGTAAACCGGCCCAACCGCGTCGCG 750156
QY 163 Glu-----ProValThrGlyHisPhePheAspGlyAspGlyMetValHisAlaVal 180
DB 750157 GAGGTGACCGCGCCACC---TACCACCTGTTTCACCGCGGACCGCATGTGTGACCGAGTTC 750213
QY 181 LysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheVal 200
DB 750214 GCGTGGCGGACGGGAAGGCC-----CGCTGGTAT 750243
QY 201 GlnGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGlu----- 216
DB 750243 -----CTGTATAACAGGACTCGTGTACCGGGTCCAGCAGC 751275

```

```

DB 750244 CGCAATCGCTGGTCCGCACACCCCGCGGTGTCGCGCGCCCTGGGCGAGCCCATTTTCGGCC 750303
QY 217 ---LeuHisGlyHisThrGlyIleAlaAlaArgLeuMetLeuPheTyrAlaArgAlaAla 235
DB 750304 CGGCTCTACCCGCGCACCGGGATTATCAG----- 750333
QY 236 GlyIleValAspProAlaHisGlyThrGlyValAlaAsnAlaGlyLeuValTyrPheAsn 255
DB 750334 -----GGCGGTCCCAACACACCGATGCTGACCCACGCC 750366
QY 256 GlyArgLeuLeuAlaMetSerGluAspAspLeuProTyrGlnValGlnIleThrProAsn 275
DB 750367 GGACGACCCCTGGCTTGTGAGCCCGCGGTG-----GTCAACTACGAATCACC 750417
QY 276 GlyAspLeuLysThrValGlyArgPheAspGlyGlnLeuGluSerThrMetIle 295
DB 750418 GATGAGCTGGACACCGTGGACCCCTGTGACTTCGACGGCACCCCTGCACGGCGTTACACC 750477
QY 296 AlaHisProLysValAspProGluSerGlyGluLeuPheAlaLeuSerTyrAspValVal 315
DB 750478 GCGCATCGCAGCGGTATCCGACACCGGTGAACCTGCACGGGTGCTCTACTCTGTTCCGCC 750537
QY 316 SerLysProTyrLeuLysTyrPheArgPheSerProAspGlyThrLysSerProAspVal 335
DB 750538 CGCGGACACAGAGTGCAGTACTCGGTGATCGCGACCGACGACGCTGCTCGGACGTT 750597
QY 336 GluIleGlnLeuAspGlnProThrMetHisAspPheAlaIleThrGluAsnPheVal 355
DB 750598 GATATCGAGTGGCGGATCCCGATGATGACAGCTTCTCCCTACCGACAACTACGTG 750657
QY 356 ValValProAspGlnGlnValPhe----- 364
DB 750658 GTGATCTACGACTCGCGGTGACCTTCAGCCCAATGACAGTGGTCCGCGCGTCCGTGCCA 750717
QY 365 -----LysLeu 366
DB 750718 CGTGGTGTCAACGCGCCCGCAGGTGTGTGATCCAGTCCGTCTGGGCGGTCTCGCATC 750777
QY 367 ProGluMetIle-----ArgGlyGlySer-----ProVal 376
DB 750778 CCGGACCGATAGCGCGGTGGGCAACCGGATGAGGGTCACTCCGATCGCTCCCGTAC 750837
QY 377 ValTyrAspLysAsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSer 396
DB 750838 GCCTGGAAACCGACGTACCTCCGCGCGTGTGTCATCGCGCGCGAGGT---GGCAAC 750894
QY 397 SerAsnIleLysTrpIleAspAlaProAspCysPheHisLeuTrpAsnAlaTrp 416
DB 750895 GAGGACGTCGCGTGTTCGACATCGAACCTCGTACGTATACCCCACTTAACCCCTAC 750954
QY 417 GluGluProGluThrAsp---GluValValIleGly-----SerCysMet 431
DB 750955 TCGGAGTCCCGAACCGCGCTGAGTGTGTGTGGAGTGGTGGCTACTACCGATG 751014
QY 432 ThrProProAspSerIlePheAsnGluSerAspGluAsnLeuLysSerValLeuSerGlu 451
DB 751015 TTTGATCCGACCGCGCGGTCCCGCGGTGACAGC-----CGGCGCTCGTGGATCGC 751068
QY 452 IleArgLeuAsnLeuLysThrGlyGluSerThrArgArgProIleIleSerAsnGluAsp 471
DB 751069 TGGACCATCAACCTGGCAGCGGTGCGGTGACCGCGAA-----TGCGCGCAGCAT 751119
QY 472 GlnGlnValAsnLeuGluAlaGlyMetValAsnArgAsnMetLeuGlyArgLysThrLys 491
DB 751120 CGGGG-----CAGGAGTTTCCCGCATCAACGAGACTCTGGTGGGTGGCGCGCATCGC 751173
QY 492 PheAlaTyr-----LeuAlaLeuAla 498
DB 751174 TTCGCTACACCGTCGGCATCGAGGGTGGTTCCTCGTCGCGCGCGCGGTGCGTGTGTCG 751233
QY 499 GluProTrpProLysValSerGlyPheAlaLysValAspLeuThrThrGlyGluValLys 518
DB 751234 ACTCG-----CTGTATAACAGGACTCGTGTACCGGGTCCAGCAGC 751275

```

```
Qy 519 LysHisLeuTyrGlyAspAsnArgTyrGlyGluProLeuPheLeuProGlyGluGly 538
Db 751276 GTCCGCTCGCTCGATCCCGACCTGCTGATCGCGGAGATGGTGTTCGTGCCGAACCCGTCG 751335
Qy 539 GlyGlu---GluAspGluGlyTyrIleLeuCysPheValHisAspGluLysThrTrpLys 557
Db 751336 CGCGGTGCAGAGATGACGGGATTCATGGCTACGGCTGGCAGCGCGCGCGAGAA 751395
Qy 558 SerGluLeuGlnIleValAsnAlaValSerLeuGluValGluAlaThrValLysLeuPro 577
Db 751396 GGCAGCTGCTCTTGTCTGGATGCCAGACTCTCGAGTCGATCGGCACCGTCGACTGCCA 751455
Qy 578 SerArgValProTyrGlyPheHisGlyThrPhe-----IleGly 590
Db 751456 CAGCGTGTGGCATGGGCTTCACGGCAACTGGGCGCGCACCTGACGGCGCTCGGG 751515
Qy 591 AlaAsp 592
Db 751516 TCGCAT 751521

RESULT 3
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.: 5,52e-37 Length: 4403765
Score: 475.50 Matches: 165
Percent Similarity: 40.35% Conservative: 90
Best Local Similarity: 26.11% Mismatches: 255
Query Match: 15.10% Indels: 123
DB: 3 Gaps: 23

US-09-758-269-6 (1-599) x US-09-103-840A-2 (1-4403765)
Qy 17 GlyAsnHisThrGlnProLeuSerSerSerGlnSerSerAspLeuSerTyrCysSer 36
Db 1018941 GCGCGCTACCCATCAACCCAGCGGAGACTGATCGGTGGCGGACATCCGG-----TCG 1018888
Qy 37 SerLeuProMetAlaSerArgValThrArgLysLeuAsnValSerSerAlaLeuHisThr 56
Db 1018887 GGGCTTCGGGTGGCGGATGTGCTCGACGCGGCCAAACAGGTGAGCGGCATTCGCCGTG 1018828
Qy 57 ProProAlaLeuHisPhePheProLysGlnSerSerAsnSerProAlaIleValLysPro 76
Db 1018827 ACTATCAGTCGAGNACCCGAGCGGTTCGGCAGCTGAATTCGGTGGCAGCAGGCCA 1018768
Qy 77 LysAlaLysGluSerAsnThrLysGlnMetAsnLeuPheGlnA-gAlaAlaAlaAla 96
Db 1018767 CCACGGTCAGTTTCG-----TTGTCAGCACCCACAC 1018738

Qy 97 LeuAspAlaAlaGluGlyPheLeuVal-SerHisGluLysLeuHisProLeuProLys-- 115
Db 1018737 GAGGATCGTGAATGGACATCAGATGTCCGAGTACTTGTTCGACCCCTCCCGAAGAC 1018678
Qy 116 -ThrAlaAspProSerValGlnIleAlaGlyAsnPheAlaProValAsnGluGlnProVa 135
Db 1018677 GACGACCACCCCTATCGACCGGTCCGTGGCGCACACAGACCCAGGAATGGATGCCGAC 1018618
Qy 135 lArgArgAsnLeuProValValGlyLysLeuProAspSerIleLysGlyValTyrValAr 155
Db 1018617 GACCTG-ACC---ACCCTGACAGCGGAAGTCCCGCGCAGCTGACGCGCATCTACCTCGG 1018562
Qy 155 qAsnGlyAlaAsnProLeuHisGluProValThrGlyHisPhePheAspGlyAspGl 175
Db 1018561 CAACCCGAGAACCCGCTACACCGCGCATTCGGACCTACCCCTTCGATGGCGACGG 1018502
Qy 175 yMetValHisAlaValLysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGl 195
Db 1018501 CATGATCCATGTCTCGCTTCGCTGATGGAAAAAGCCTTCTACCGCAACCGATTTATTCG 1018442
Qy 195 nThrAsnArgPheValGlnArgGlnLeuGlyArgProValPheProLysAlaIleGl 215
Db 1018441 CACCGATGGATTCTTGCCGAGAACGAGCGCGCGCGCTGTGGCGC----- 1018393
Qy 215 yGluLeuHisGlyHisThrGlyIleAlaArgLeuMetLeuPheTyrAlaArg----- 232
Db 1018392 -----GGTCTGGCAGAACCGGTGCAACTGGCCAGCGGGGAACACGG 1018352
Qy 233 ----AlaAlaAlaGlyIleValAspProAlaHisGlyThrGlyVal---AlaAsnAlaGl 250
Db 1018351 CTGGGCGCTCGTGGCCTCATGAAGGACGCGTGAGCACCAGCGCTCATCGTCCACCGAGG 1018292
Qy 250 yLeuVal-----TyrPheAsnGlyArgLeuLeuAlaMetSerGluAspLe 266
Db 1018291 TATCGCGCTGACTAGCTTCTACCGAGTGGCGCATCTGTGCGATCGAC----- 1018243
Qy 266 uProTyrGlnValGlnIleThrProAsnGlyAspLeuLysThrValGlyArgPheAspPh 286
Db 1018242 -CGTACTCGGCCAATACGGTC-----GGCAGGAGAGTGGCAGCGAAGGTTTCGTT 1018190
Qy 286 eAspGlyGlnLeuGluSerThrMetIleAlaHisProLysValAspProGluSerGlyGl 306
Db 1018189 CGACTGGCGCGTGTGCG-----GCACATCCGAAGGTAGACAAACAGACCGCGCA 1018142
Qy 306 uLeuPheAlaLeuSerTyrAspValValSerLysProTyrLeuLysTyrPheArgPheSe 326
Db 1018141 ACTGTTGTTCTCAACTACAGC---AAGCAAGAGCCGTATATGCGCTACCGCGTTGTCGA 1018085
Qy 326 rProAspGlyThrLysSerProAspValGluIleGlnLeuAspGlnProThrMetMetHi 346
Db 1018084 CCAGAACAAATGAGCTTGTGCACTATGTCGATGTTCGCTGCCCGGCGCGGTACCGCA 1018025
Qy 346 sAspPheAlaIleThrGluAsnPheValValValProAspGlnGlnValValPheLysLe 366
Db 1018024 TGACATGGCGTTCACCGAAAAATTACGTAATCCTCAACGAT----- 1017985
Qy 366 uProGluMetIleArgGlySerProValValTyrAspLysAsnLysVal----- 383
Db 1017984 -----TTTCCACTGTTTCGGGATCCCGAGGTGCTCGAGCGCGCA 1017947
Qy 384 -----AlaArgPheGlyIleLeuAspLysTy 392
Db 1017946 GTGCACCTACCAGCTTCTATCCGGAGATTCATCTCGGTTTCGCGTGT----- 1017895
Qy 392 rAlaGluAspSerSerAsnIleLysTrpIleAspAlaProAspCysPheCysPheHisLe 412
Db 1017894 -GCTCGCGGAGGCAACAGCATTCGCTGTTCGAAGCCGATCCAACGTTTCGTTGCTACTT 1017836
Qy 412 uTrpAsnAlaTrpGluGluProGluThrAspGluValValIleGlySerCysMetTh 432
Db 1017835 CACCAACGCTTACGAG-----CAGGCGCAGAGATCGTCTCGACGCG----- 1017793
Qy 432 rProAspSerIlePheAsnGluSerAsp----- 442
```

```

Db 1017792 -----TTCTACGAGCGCATCGCAGCCACTTGCACCGAGGAAACGAA 1017749
QY 443 -----GluAsnLeuLysSerValLeuSerGlu1 452
Db 1017748 GTGGGAGAGCTGTTTCGGTTCCTGCTCTGGATCGGTGCTACCTCCGGCTACATCGGTG 1017689
QY 452 eArgLeuAsnLeuLysThrGlyGluSerThrArgArgProIleIleSerAsnGluAspG1 472
Db 1017688 GCGGCTCAATATGTTGACCGGG-----GCAGTCCACGAGGA 1017653
QY 472 nGlnValAsn-----LeuGluAlaGlyMetValAsnArgAsnMetLeuGlyArgly 489
Db 1017652 GCACCTGTCGAGTCATCACCGAGTTCGGAACCATCAACCGGATTCACGCGCCAGCAG 1017593
QY 489 sThrLysPheAlaTyrLeuAlaLeuAlaGluPro---TrpProLysValSerGlyPheAl 508
Db 1017592 CTACCGTTACACTATGCGCTACCGGCAACCGAGTTGTTCTTCTTCGACGGACTGCT 1017533
QY 508 aLysValAspLeuThrThrGlyGluValLysLysHisLeuTyrGlyAspAsnArgTyrG1 528
Db 1017532 TAAGCACGATCTGCTCACCGGAACACACAGAGTGTACTGTTCTGCTGACGGCTCTACGG 1017473
QY 528 yGlyGluProLeuPheLeuProGlyGluGlyGlu---GluAspGluGlyTyrIleLe 547
Db 1017472 AGTGAGACCGGATGGCTCCACGGGTGGGCGAGCAGCGCCGAGGACGAGCGTATCTGCT 1017413
QY 547 uCysPheValHisAspGluLysThrTriPlysSerGluLeuGlnIleValAsnAlaValSe 567
Db 1017412 CACCCTCACCGACATGAACGACGACGATCTGATTGCTTGGTTTCGACCGCGCGG 1017353
QY 567 rLeuGluValGluAlaThrValLys-----LeuProSerArgValProTyrGlyPheHi 585
Db 1017352 CCCCAGCGATGCGCCGATATGCAAGCTTGCACTGCCGGAACGTATTTCCAGCGGCACGCA 1017293
QY 585 sGlyThrPheIleGlyAlaAspLeuAlaLys 596
Db 1017292 TTCGGCGTGGTGCSCGGCGCGCGAGTTCGTCGC 1017259

RESULT 4
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37rv
US-09-103-840A-1

Alignment Scores:
Pred. No.: 5,53e-37 Length: 4411529
Score: 475.50 Matches: 165
Percent Similarity: 40.3% Conservative: 90
Best Local Similarity: 26.11% Mismatches: 255
Query Match: 15.10% Indels: 123
DB: 3 Gaps: 23

US-09-758-269-6 (1-599) x US-09-103-840A-1 (1-4411529)

QY 17 GlyAsnHisThrGlnProLeuSerSerGlnSerSerAspLeuSerTyrCysSer 36
```



16390	Db	-----GAACTGGATCTCTG---ACACTTGAACATCGCGGACACTTTCGACTACGACGGC	16340
289	Qy	GlnLeuGluSer---ThrMetIleAlaHisProLysValAspProGluSerGlyGluLeu	307
16339	Db	CAAGTTACCGCAACCCACACCGCCATCCAAAATATGACCCGAAACGGGTGACITG	16280
308	Qy	PheAlaLeuSerTyrAspValValSerLysProTyrLeuLysTyrPheArgPheSerPro	327
16279	Db	TTG-----TTCCTTCGGTTCGCAGCT	16259
328	Qy	AspGlyThrLysSerProAspVal-----	335
16258	Db	AAGGGCAAGCAACTCAGACATGGCTATTACATTGTCGACAGCAGCGCAAGGTGACA	16199
336	Qy	---GluIleGlnLeuAspGlnPro-----ThrMetMetHisAspPheAlaIleThrGlu	352
16198	Db	CATGAAACTTTGGTTGAGCAGCCCTATGGCGCATTCATGCACGACTTTGCCATTACCCGA	16139
353	Qy	AsnPheValValProAspGlnGlnValPheLysLeuProGluMetIleArgGly	372
16138	Db	AAITGGTGCCATTTCCCAATTATTCGCGGCGCCACCAACAGCCTG---TCCCGCTCAAGGCG	16082
373	Qy	GlySerProVal---ValTyrAspLysAsnLysValAlaArgPheGlyIleLeuAspLys	391
16081	Db	AAACAGCAATTTATATTGGGGAGCCGGAACCTGCGGCGAGCTACATTGGCGGTACTC	16028
392	Qy	TyrAlaGluAspSerSerAsnIleLysTrpIleAspAlaProAspCysPheCysPheHis	411
16027	Db	CGCGCGCCAGGCAGTCTGATTTCCTGGCTCMAGGCACCGCGCTCTGGGTATTTCAT	15968
412	Qy	LeuTrpAsnAlaTrpGluGluProGluThrAspGluValValIleGlySerCysMet	431
15967	Db	GTGTGAAATCTTTGGGAAGTCGGAACCAAGATTATATCGACTTATGGAAGTGAATC	15908
432	Qy	ThrPro-----ProAspSer-----	436
15907	Db	CTCGCGTTCCTTCCTCCCAACTCACAACCAACCCCTTCGCCCCCTGAGAAAGCCGTACCA	15848
437	Qy	-----IlePheAsnGluSerAspGluAsnLeuLysSerVal	448
15847	Db	CGCCTGACTGTTGGGAAATTGACCTCGATAGCAGCGACGACAGATCAAGGGAACCCGG	15788
449	Qy	LeuSerGluIleArgLeuAsnLeuLysThrGlyGluSerThr-----	462
15787	Db	CTACAGATTCTTTGGGAAATGCCAATCATGGATTCTTCGTCGCCCTGCAATGCAAC	15728
463	Qy	-----ArgArgProIleIleSerAsnGluAspGln	472
15727	Db	CGTATGCTGTTTATGGGGGTGGACGATCCACCAACCACTTTCGCATCAGCAGCCGAG	15668
473	Qy	GlnValAsnLeuGluAlaGlyMetValAsnArgAsnMetLeuGlyArgLysThrLysPhe	492
15667	Db	AAAGATA-----TTC	15659
493	Qy	AlaTyrLeuAlaLeuAlaGluProTrpProLysValSerGlyPheAlaLysValAspLeu	512
15658	Db	GCGTACAACTCACTCGC---ATCTGG-----GACAAC	15629
513	Qy	ThrThrGlyGluValLysLysHisLeuTyrGlyAspAsnArgTyrGlyGlyGluProLeu	532
15628	Db	CACCGAGGTGACTACGACCTCTGGTACTCCGCGGAAGCCCTCGCGCGGCCAGGAGCCGCC	15569
533	Qy	PheLeuPro---GlyGluGlyGlyGluAspGluGlyTyrIleLeuCysPheValHis	551
15568	Db	TTCGTCCCTAGAAAGTCCGACCGCGCCGAAGGTGACGGGTACTCTGCTACCGTGTGGT	15509
552	Qy	AspGluLysThrTrpLysSerGluLeuGlnIleValAsnAlaValSerLeuGlu-----	569
15508	Db	CGCCTCGATGAAATCGCAGCGATCTGGTAAATTCGACACTCAGACATCCAGCTCGGT	15449
570	Qy	ValGluAlaThrValLysLeuProSerArgValProTyrGlyPheHisGlyThrPheIle	589
15448	Db	CCCGTGGCAACCATCAAGCTGCTCATTCGGGTAAAGGGCGCGCTCTCATGGCTGTGGGTA	15389

```

RESULT 6
US-08-976-063E-21/c
; Sequence 21, Application US/08976063E
; Patent No. 6524831
; GENERAL INFORMATION:
; APPLICANT: Steinhüchel, Alexander
; APPROPRIATOR: Priefert, Horst
; INVENTOR: Rabenhöft, Jürgen
; TITLE OF INVENTION: SYNTHETIC ENZYMES FOR THE PRODUCTION OF CONIFERYL ALCOHOL, CONFERYLALEHYDE, FERULIC ACID, VANILLIN AND DERIVATIVES THEREOF
; FILE REFERENCE: Beyer-9998-CAO
; CURRENT APPLICATION NUMBER: US/08/976,063E
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 196 49 655.1 GERMANY
; FIRST FILING DATE: 1996-11-29
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 1518
; TYPE: DNA
; ORGANISM: not required under old rule
US-08-976-063E-21

Alignment Scores:
Pred. No.:          9,53e-36           Length:      1518
Score:              411.50             Matches:     137
Percent Similarity: 40.68%             Conservative: 79
Best Local Similarity: 25.80%            Mismatches:   192
Query Match:        13.06%             Indels:       123
DB:                  4                 Gaps:         19

US-09-758-269-6 (1-599) x US-08-976-063E-21 (1-1518)

Qy    118 AspProSerValGlnIleAlaGlyAsnPhenPheAlaProValAsnGluGlnProValArgAsg 137
Db    1497 GACCGC-----CAATTAGTAGGAACAACCTTCTCCACCAGGTATAGAGGCAGACTTGTTTC 1444

Qy    138 AsnLeuproValValgLylysLeuProAspSeriLeysGlyValTyrrValArgAsnGly 157
Db    1443 GATCTAGAGTTTGACGGCGAAATCCCCAAAATCAATAAATGGAACGTTTTCTACCGTAATACG 1384

Qy    158 AlaasnProLeuhisGluProValThrGlyHisHisPhePheAspGlyMetVal 177
Db    1383 CGAGAGCTCAAGTTACCCCACAAAAATCCACACCTTCATAGATGAGATGAATGGGCC 1324

Qy    178 HisAlaValylsPheGluHisGlyserAlaserTyrAlaCyseArgPheThrGlnThrAsn 197
Db    1323 TCTGCCTTCACCTTCGAAGATGCTCATGTGCGACTTCATCAGTCGCTGGGTAAACACCGCT 1264

Qy    198 ArgPhevalGlnGluargGlnLeuGlyArgProValPheProlysAlaleGlyGluLeu 217
Db    1263 CgatTCACGGCCGAACGACTAGCGCAAATCGCTATTTGGCATGTACAGAACCCTAT 1204

Qy    218 HissglyHisThrglyileAlaaArgLeumetLeuPheTyrAlaaArgAlalaAlaGlylle 237
Db    1203 ACCGACGACACCACTGTAAGAGACTA----- 1177

Qy    238 ValaspProAlahisglYthrGlylValalaaSnAlaglyLeuValtyrPheasnglyArz 257
Db    1176 -----CACCGCACCGTGGCCAATACAAGCATCATTAGCCATCACGGCAAG 1132

Qy    258 LeuleAlametserCluahspaspleuProtyrGlnvalGlNleThrpAsnGlyAsp 277
Db    1131 GTGCTCGCGGTGAAGAACGAGCGCCCTACCGTAC-----GAATGGATCCTCGT---ACA 1081

Qy    278 LeulystrValclYArgPheaspPheaspGlyGinLeuglusEr----ThrmellAlela 296
Db    1080 CTTGAAACTCGGGACACTTCGNACTACGACGGCCAAGTTACCAGGCCAAACCCACACCGCC 1021

Qy    297 HisprolylsValaspProgluserglyGluLeuPheAlaLeuSerTyrAspValValSer 316

```



```
Db 1020 CATCAAAATATGACCGGAAACGGGTGACTTGTTG----- 985
Qy 317 LysProTyrLeuLysTyrPheArgPheSerProAspGlyThrLysSerProAspVal--- 335
Db 984 -----TTCTTCGGTTCGGCAGCTAAGGCGGAAGCAACTCCACAGATGCC 940
Qy 336 -----GluIleGlnLeuAspGlnPro--- 342
Db 939 TATTACATTGCGACACGCGCAAGGTGACATGAACCTTGGTTTGACGACCCCTAT 880
Qy 343 ---ThrMetMetHisAspPheAlaIleThrGluAsnPheValValProAspGlnGln 361
Db 879 GCGCATTCATGACGACATTTTCCATTACCGAAATTCGTCCATTTTCCCAATTATGCCG 820
Qy 362 ValValPheLysLeuProGluMetIleArgGlyGlySerProVal---ValTyrAspLys 380
Db 819 GCCACCAACAGCCTG---TCCCGCTCAAGCGCAACACGCAATTTATATGTGGAGCGG 763
Qy 381 AsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSerSerAsnIleLys 400
Db 762 GAATGGCGCATACATTGGCGTACTC-----GGCGCGCCGACGAGTCTGATTCCG 709
Qy 401 TrpIleAspAlaProAspCysPheCysPheHisLeuTrpAsnAlaTrpGluGluProGlu 420
Db 708 TGGCTCAAGGCACCGCGCTCTGGTATTTTCATGTTGTGAATGCTTGGGAAGTCGGAACC 649
Qy 421 ThrAspGluValValIleGlySerCysMetThrPro-----ProAspSer--- 436
Db 648 AAGATTATATCGACCTTATGAAAGTGAATCTCGCGTTTCCCTTTCGCCCAACTCAAA 589
Qy 437 -----Ile 437
Db 588 AACCAACCTTCGCGCCCTGAGAAAGCCGTACACGCGCTGACTCGTTGGGAAATGACCTC 529
Qy 438 PheAsnGluSerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsnLeuLys 457
Db 528 GATAGCAGCAGCAGAGATCAAGCGAACCGGCTACACGATTTCTTTGCGGAATGCCA 469
Qy 458 ThrGlyGluSerThr----- 462
Db 468 ATCATGGATTCTTCGTCGCCCTGCAATGCAACCGCTATGGCTTTATGGGGTGGACGAT 409
Qy 463 ---ArgArgProIleIleSerAsnGluAspGlnGlnValAsnLeuGluAlaGlyMetVal 481
Db 408 CCACGCCAAACCACTTGCCTCATCAGCAGCCCGAGAGATA----- 370
Qy 482 AsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeuAlaLeuAlaGluProTrp 501
Db 369 -----TTGCGGTACAACCTCAGCTGGC---ATCTGG 343
Qy 502 ProLysValSerGlyPheAlaLysValAspLeuThrThrGlyGluValLysLysHisLeu 521
Db 342 -----GACAACCAACCGAGGTGACTACGACCTCTGGTAC 310
Qy 522 TyrGlyAspAsnArgTyrGlyGlyGluProLeuPheLeuPro---GlyGluGlyGlyGlu 540
Db 309 TCCGGCGAAGCTCGCGCGCCAGAGCGCGCTTCGTCCTTAGAAGTCCGACCGCGCCG 250
Qy 541 GluAspGluGlyTyrIleLeuCysPheValHisAspGluLysThrTrpLysSerGluLeu 560
Db 249 GAAGGTGACGGGTACTTCTGCTACCGTGGTGTGGTCCCTCGATGAAATCGCAGCGATCTG 190
Qy 561 GlnIleValAsnAlaValSerLeuGlu-----ValGluAlaThrValLysLeuProSer 578
Db 189 GTAATTCCTGACACTCAAGACATCCAGTCTGGTCCCGTGGCAACCATCAAGCTGCCATC 130
Qy 579 ArgValProTyrGlyPheHisGlyThrPheIle 589
Db 129 CGGCTAAGGCGCGCTCTCCATGGCTGCTGGTA 97
```

RESULT 7

US-08-200-807-1

; Sequence 1, Application US/08200807

```
; Patent No. 5573939
; GENERAL INFORMATION:
; APPLICANT: B vdk, Claes Olof, Eriksson, Ulf
; TITLE OF INVENTION: Isolated Protein Receptors, Antibodies Which
; TITLE OF INVENTION: bind Thereeto, Nucleic Acid Sequence Coding
; Patent No. 5573939
; TITLE OF INVENTION: Therefor, And Uses Thereof
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA: US/08/200,807
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/883,539
; FILING DATE: 15-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5573939man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 280
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2629 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHEICAL: no
; ANTI-SENSE: no
; US-08-200-807-1
;
; Alignment Scores:
; Pred. No.: 5,23e-13 Length: 2629
; Score: 211.00 Matches: 130
; Percent Similarity: 40.03% Conservative: 99
; Best Local Similarity: 22.73% Mismatches: 215
; Query Match: 6.70% Indels: 131
; DB: 1 Gaps: 26
;
; US-09-758-269-6 (1-599) x US-08-200-807-1 (1-2629)
;
; Qy 105 ValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSerValGlnIleAla 124
; Db 14 ATGTCCAGCAAGTGAACATCCA-----GCT 40
;
; Qy 125 GlyAsnPheAlaProValAsnGluGlnProValArgArgAsnLeuPro----- 140
; Db 41 GGTGTTTACAAGAACTGTTTGAACCTGTGGAGAACTATCTCCACCGCTCACAGCCAT 100
;
; Qy 141 ValValGlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyAlaAsnPro 160
; Db 101 GTTACAGCGAGATCCCTCTGGCTAACCGCGAGTCTCTCTTCGATGTGGGCCAGGACTC 160
;
; Qy 161 LeuHis-----GluProValThrGlyHisHisPheAspGlyAspGlyMetVal 177
; Db 161 TTGAGGTGGATCGGAACCATTT-----TACCACCTGTTTGAATGGCAAGCCCTCTCTA 214
;
; Qy 178 HisAlaValLysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsn 197
; Db 178 HisAlaValLysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsn 197
```

Db 215 CACAGTTTGGACTTTAAAGAGGACATGTCAATACACAGAGGTTTCATCCGCACTGAT 274  
 Qy 198 ArgPheValGlnGlu-----ArgGlnLeuGlyArgProVal 209  
 Db 275 GCTTACGTACGGGCAATGACTGAGAAAGGATCGTCATAACAGAAATTTGGCACCCTGTGCT 334  
 Qy 210 PheProLysAlaIleGlyLeuHisGlyHisThrGlyLeuAlaArgLeuMetLeuPhe 229  
 Db 335 TTCCCGAGATCCCTGCAAGAAATATA-----TTTCCAGGTTTTTTTCTTAC 379  
 Qy 230 TyrAlaArgAlaAlaAlaGlyIleValAspProAlaHisGlyThrGlyValAlaAlaAla 249  
 Db 380 TTC-----CGAGGAGTGGAGGTACTGACAT 406  
 Qy 250 GlyLeuValTyrPhe-----AsnGlyArgLeuLeuAlaMetSerGluAspLeu 266  
 Db 407 GCCCTTGTGTTAA-TATCTACCCAGTGGGGAAGATTACTATGCTGCACAGACCAACTT 465  
 Qy 267 ProTyr-GlnValGlnIleThrProAsnGlyAspLeuLysThrValGlyArgPheAsp 286  
 Db 466 CATTACAAAGGTT-----AATCCTGAG--ACCTTGGAAACAATTAAGCAGGTTGACCT 516  
 Qy 286 eAspGlyGlnLeuGlu-----SerThrMetIleAlaHisProLysValAspProGluSer-- 304  
 Db 517 TTGCAACTATGTCTAGTCAATGGAGCCACTGCTCACCCACATTTGAAATGATGGGAC 576  
 Qy 305 -----GlyGluLeuPheAlaLeuSerTyrAspValIse 316  
 Db 577 TGTTTACCAATTTGGTAATGCTTTGGGAAAAATTTTCAATTCCTACAAATTTGTA 636  
 Qy 316 rlyAsProTyrLeuLysTyrPheArgPheSerProAspGlyThrLysSerProAspVal 336  
 Db 637 GATCCCACTACCTACCAAGCAGACAGCAAGATGCA-----ATAAGCAAGTCAGAGATCGT 690  
 Qy 336 uileGlnLeuAsp-----GlnProThrMetMetHisAspPheAlaIleTh 351  
 Db 691 TGTACAAATCCCTGCAGTGCAGTCAAGCCATCTACGTCATATGTTTGGTTTGC 750  
 Qy 351 rGluAsnPheValValProAspGlnValValPheLysLeuProGluMetIleAr 371  
 Db 751 TCCCAACTATATGTTTGTGGAGACACCACTCAAAATTAATCTGTTCAAGTTCTTTC 810  
 Qy 371 gGlyGlySerProValValTyrAspLysAsnLysValAlaArgPhe----- 386  
 Db 811 TTCTGAGT-----CTTGGGAGCCAAATACATGATGTTTGTGAATCCAATGAAC 864  
 Qy 387 -----GlyIleLeuAspLysTyrAlaGluAspSerSerAsnIleLysTr 401  
 Db 865 CATGGGGTGTGCTTTCATATGCTGACAAAAAAGAAAGTATATCAATAATAATA 924  
 Qy 401 pIleAspAlaProAspCysPheCysPheHisLeuTrpAsnAlaTrpGluGluProGluTh 421  
 Db 925 CAGGACCTCTCT-----TTTAACCTCTTTCATCATCAATCACTATGAAGCATGAT 981  
 Qy 421 rAspGluValValValIleGlySerCysMetThrProProAspSerIlePheAsn----- 439  
 Db 982 T-----CTGATGTGGATCTCTGTTGCTGGAAGGATTTGAATTTGTTATAATATT 1035  
 Qy 440 -----GluSerAspGluAsnLeuLysSerValLeuSerGluIleAr 453  
 Db 1036 ATATTAGCAATTTACGTGAACTGGGAAGAGGTGAAAAA--ATGCCCAAGGCT 1093  
 Qy 453 gLeuAsnLeuLysThrGlyGluSer-----ThrArgAr 464  
 Db 1094 CTTACGCTGAATTTAGGATACGTACTCTTTTGAATATTGACAAAGCTGACACAGC 1153  
 Qy 464 gProIleLeuSerAsnGluAspGlnValAsnLeuGluAlaGlyMetValAsnArgAs 484  
 Db 1154 AAGAATTTAGTCACACTCCCAACACAACTGCCCTCAATCTCTGTCAGTGACGAGACC 1213  
 Qy 484 mMetLeuGlyArgLysThrLysPheAlaTyrLeuAlaLeuAlaGluProTrpProLysVa 504  
 Db 1214 ATCT---GGCTGGAACCTGAGGTCTCTTTTCAGGGCCTCGCCCAAGCATTTGAGTTTCT 1270

Qy 504 IserGlyPheAlaLys-----ValAspLeuThr-- 513  
 Db 1271 CAATCAATTAACAGAGATATGTTGGGAAACCTTACACATATGCATATGACCTTGGCTTG 1330  
 Qy 514 -----ThrGlyGluValLysHisLeuTy 522  
 Db 1331 AATCACTTTTCCAGACAGGCTCTGTAAAGCTGAACGTCAAAACTAAAGAAACCTGGGTA 1390  
 Qy 522 rGlyAsp-----AsnArgTyrGlyGlyGluProLeuPheLeuPro---GlyGluGlyG 539  
 Db 1391 TGGCAA-CAGCCTGATTCATACCCCTCAGAACCTATCTTTTCTCACCAGATGCCTT 1449  
 Qy 539 yGluGluAspGluGlyTyrIleLeuCysPheValHisAspGluLysThrTrpLys----- 557  
 Db 1450 GGAGGAAGATGACGCTGTAGTCTGTAGTGTGTGAGCCCTGGGGCAGGACAAAAGCC 1509  
 Qy 558 -SerGluLeuGlnIleValAsnAlaValSerLeuGluValGluAlaThrValLysLeuPr 577  
 Db 1510 TGCTTATCTTCTGATTTCTGAATGCCAAGGACTTGAGTGAAGTTGCCAGGCTGAAGTGA 1569  
 Qy 577 oSerArgValProTyrGlyPheHisGlyThrPhe 588  
 Db 1570 GATTAACATCCCTGACCTTTCATGGACTGTT 1603

## RESULT 8

US-08-488-305A-1  
 ; Sequence 1, Application US/08488305A  
 ; Patent No. 5679772  
 ; GENERAL INFORMATION:  
 ; APPLICANT: B vik, Claes Olof; Eriksson, Ulf; Peterson, Per A.  
 ; TITLE OF INVENTION: Isolated Protein Receptors, Antibodies Which  
 ; TITLE OF INVENTION: bind Thereto, Nucleic Acid Sequence Coding  
 ; Patent No. 5679772  
 ; TITLE OF INVENTION: Therefor, And Uses Thereof  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Felfe & Lynch  
 ; STREET: 805 Third Avenue  
 ; CITY: New York City  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10022  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 kb storage  
 ; COMPUTER: IBM PS/2  
 ; OPERATING SYSTEM: PC-DOS  
 ; SOFTWARE: Wordperfect  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/488,305A  
 ; FILING DATE: 7-JUNE-1995  
 ; CLASSIFICATION: 530  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Kohli, Vineet  
 ; REGISTRATION NUMBER: 37,003  
 ; REFERENCE/DOCKET NUMBER: LUD 5280.3  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 688-9200  
 ; TELEFAX: (212) 838-3884  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2629 bases  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA to mRNA  
 ; HYPOTHEITICAL: no  
 ; ANTI-SENSE: no  
 ; US-08-488-305A-1

Alignment Scores: 5.23e-13 Length: 2629  
 Pred. No.: 211.00 Matches: 130  
 Score:

Percent Similarity: 40.03% Conservative: 99  
Best Local Similarity: 22.73% Mismatches: 215  
Query Match: 6.70% Indels: 131  
DB: 1 Gaps: 26

US-09-758-269-6 (1-599) x US-08-488-305A-1 (1-2629)

Qy 105 ValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSerValGlnIleAla 124  
Db : : : : :  
Qy 14 ATGTCCAGCCCAAGTTGAACATCCA-----GCT 40  
Db : : : : :  
Qy 125 GlyAsnPheAlaProValAsnGlnProValArgAsnLeuPro----- 140  
Db : : : : :  
Qy 41 GGTGGTTACAGAAACTGTTTGAACCTGTGGAGNACTATCTCCACCGCTCACAGCCCAT 100  
Db : : : : :  
Qy 141 ValValGlyLysLeuProAspSerIleLysGlyValTyValArgAsnGlyAlaAsnPro 160  
Db : : : : :  
Qy 101 GTTACAGCAGGATCCCTCGCTAACCCGCACTCTCTTCGATGTGGGCCAGGACTC 160  
Qy 161 LeuHis-----GluProValThrClyHisHisPheAspGlyAspGlyMetVal 177  
Db : : : : :  
Qy 161 TTTGAGTTGATCGGAACCAATTT-----TACCACCTGTTTGTGGCAAGCCCTCCTA 214  
Db : : : : :  
Qy 178 HisAlaValLysPheGluHisGlySerAlaSerTyValAlaCysArgPheThrGlnThrAsn 197  
Db : : : : :  
Qy 215 CACAAGTTTGACTTTAAAGAGGACATGTCATACACACAGAGTTTCATCCGCACTGAT 274  
Qy 198 ArgPheValGlnGlu-----ArgGlnLeuGlyArgProVal 209  
Db : : : : :  
Qy 275 GCTTACGTACGGCAATGACTGAGAAAAGGATCGTCATAACAGAAATTTGGCACCTGTGCT 334  
Qy 210 PheProLysAlaIleGlyGluLeuHisGlyHisThrGlyIleAlaArgLeuMetLeuPhe 229  
Db : : : : :  
Qy 335 TTCCAGATCCCTGCAAGAAATA-----TTTCCAGGTTTTTTTCTTATC 379  
Qy 230 TyValAlaArgAlaAlaGlyIleValAspProAlaHisGlyThrGlyValAlaAsnAla 249  
Db : : : : :  
Qy 380 TTC-----CGAGGAGTGGAGTTTACTGACAAAT 406  
Qy 250 GlyLeuValTyPhe-----AsnGlyArgLeuLeuAlaMetSerGluAspLeu 266  
Db : : : : :  
Qy 407 GCCCTTGTTAA-TATCTACCCAGTGGGGAAGATTACTATGCTGCACAGAGACCAACTT 465  
Qy 267 ProTyrr-GlnValGlnIleThrProAsnGlyAspLeuLysThrValGlyArgPheAspPh 286  
Db : : : : :  
Qy 466 CATTACAAAGGTT-----AATCTGAG-----ACCTTGGAAAACAATTAGCAGGTGACCT 516  
Qy 286 eAspGlyGlnLeuGlu-----SerThrMetIleAlaHisProLysValAspProGluSer-- 304  
Db : : : : :  
Qy 517 TTGCAACTATGTCAGTCAATGAGCCATGCTCACCACCCACATTTGAAAATGATGGAC 576  
Qy 305 -----GlyGluLeuPheAlaLeuSerTyAspValValSe 316  
Db : : : : :  
Qy 577 TGTTTACAACATTTGTAATGCTTTGGGAAAAATTTTCAATTGCTTACAATAATTGTA 636  
Qy 316 rLysProTyrrLeuLysTyrrPheArgPheSerProAspGlyThrLysSerProAspValG 336  
Db : : : : :  
Qy 637 GATCCCACTTACAAGCAGACAGCAAGAGATCCA-----ATAAGCAAGTCAGAGATCGT 690  
Qy 336 uileGlnLeuAsp-----GlnProThrMetMetHisAspPheAlaIleTh 351  
Db : : : : :  
Qy 691 TGTACAATTTCCCTGACGTGACCGATTTCAAGCCATCTTACGTCATAGTTTGGTTGAC 750  
Qy 351 rGluAsnPheValValValProAspGlnGlnValValPheLysLeuProGluMetIleAr 371  
Db : : : : :  
Qy 751 TCCCAACTATATGTTTGTGGACACACAGTCAAAATTAATCTGTTCAAGTTTCTTTC 810  
Qy 371 gGlySerProValValTyrrAspLysAsnLysValAlaArgPhe----- 386  
Db : : : : :  
Qy 811 TTATGAGAGT-----CTTTGGGAGCAATTACATGATGATTTTGAATCCAATGAAC 864  
Qy 387 -----GlyIleLeuAspLysTyrrAlaGluAspSerSerAsnIleLysTr 401  
Db : : : : :  
Db 865 CATGGGGTTTGGCTTCATATTGCTGACAAAAAAGAAAAAGTATATCAATAATAATA 924  
Qy 401 pIleAspAlaProAspCysPheCysPheHisLeuTrpAsnAlaTrpGluGluProGluTh 421  
Db : : : : :  
Qy 925 CAGGACCTCTCTCT--TTTAACTCTTTTCAATCATCAATACCTATATGAAGACCATGATT 981  
Qy 421 rAspGluValValIleGlySerCysMetThrProAspSerIlePheAsn----- 439  
Db : : : : :  
Qy 982 T-----CTGATTGTGGATCTCTGTGCTGAAAGGATTTGAATTTCTTTATATATTT 1035  
Qy 440 -----GluSerAspGluAsnLeuLysSerValLeuSerGluIleAr 453  
Db : : : : :  
Qy 1036 ATATTGACCAATTACCTGAGAAGCTGGGAAGAGGTGAAAAA--ATGCCAGAAAGCT 1093  
Qy 453 gLeuAsnLeuLysThrGlyGluSer-----ThrArgAr 464  
Db : : : : :  
Qy 1094 CCTCAGCTCAAGTTAGGATAGATACGTATCTTCTTGAATATTGACAAGGCTGACACAGC 1153  
Qy 464 gProIleIleSerAsnGluAspGlnGlnValAsnLeuGluAlaGlyMetValAsnArgAs 484  
Db : : : : :  
Qy 1154 AAGNATTTAGTACACTCCCAACACACTGCCACTGCAATTTCTGTCAGTGCAGACACC 1213  
Qy 484 nMetLeuGlyArgLysThrLysPheAlaTyrrLeuAlaLeuAlaGluProTrpProLysVa 504  
Db : : : : :  
Qy 1214 ATCT--GGCTGGAACTGAGGTTCTCTTTTCAGGGCTCGCCAAAGCATTTGAGTTTCT 1270  
Qy 504 lSerGlyPheAlaLys-----ValAspLeuThr-- 513  
Db : : : : :  
Qy 1271 CAATCAATTTACCAGAAATATGTTGGGAAACCTTTACATATATGATATGACTTTGGTTG 1330  
Qy 514 -----ThrGlyGluValLysLysHisLeuTy 522  
Db : : : : :  
Qy 1331 AATCACTTTGTCAGACAGGCTCTGTAACTGAAGTCAAACTAAAGAAACCTGGTA 1390  
Qy 522 rGlyAsp-----AsnArgTyrrGlyGluProLeuPheLeuPro---GlyGluGlyG 539  
Db : : : : :  
Qy 1391 TGGCAA-GAGCCTGATTACATACCCCTCAGAACCTATCTTTGTTCTCACCAGATGCTT 1449  
Qy 539 yGluGluAspGlyTyrrIleLeuCysPheValHisAspGluLysThrTrpLys----- 557  
Db : : : : :  
Qy 1450 GGAGGAAGATGACGGTGTAGTTCTGAGTGTGGTGTGAGCCCTGGGCGCAGACAAAGACC 1509  
Qy 558 -SerGluLeuGlnIleValAsnAlaValSerLeuGluValGluAlaThrValLysLeuPr 577  
Db : : : : :  
Qy 1510 TGCTATCTTCTGATTCGATGCAAGGACTTGAGTGAAGTTGCCAGGGCTGAAGTGA 1569  
Qy 577 oSerArgValProTyrrGlyPheHisGlyThrPhe 588  
Db : : : : :  
Qy 1570 GATTAACATCCCTGCTCACCCTTTTCATGACTGTTT 1603

RESULT 9  
US-09-385-259-1  
; Sequence 1, Application US/09385259  
; Patent No. 6201114  
; GENERAL INFORMATION:  
; APPLICANT: Aguirre, Gustavo D.  
; APPLICANT: Acland, Gregory M.  
; APPLICANT: Ray, Kunal  
; TITLE OF INVENTION: IDENTIFICATION OF CONGENITAL STATIONARY NIGHT BLINDNESS  
; TITLE OF INVENTION: IN DOGS  
; FILE REFERENCE: 19603/2481  
; CURRENT APPLICATION NUMBER: US/09/385.259  
; CURRENT FILING DATE: 1999-08-30  
; EARLIER APPLICATION NUMBER: 60/103,219  
; EARLIER FILING DATE: 1998-10-06  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 1724  
; TYPE: DNA  
; ORGANISM: Canis familiaris  
US-09-385-259-1

Alignment Scores:

Assembly Coordinates:					
Pred. No.:	3,3e-13	Length:	1724		
Score:	210.00	Matches:	113		
Percent Similarity:	37.08%	Conservative:	98		
Best Local Similarity:	19.86%	Mismatches:	218		
Query Match:	6.67%	Indels:	140		
DB:	3	Gaps:	25		
US-09-758-269-6 (1-599) x US-09-385-259-1 (1-1724)					
Qy	120	SerValGlnIle-----AlaGlyAsnPheAlaProValAsnGluGlnProValArg	136		
Db	31	TCCATCCAAGTGGAGCATCCCGCGCGGTTCACAGAAGCTGTTTGAACCGTGGAAGAG	90		
Qy	137	ArgAsnLeuPro-----ValValGlyLysLeuProAspSerIleLysGlyVal	152		
Db	91	CTGCTGTCGCCGTCAACC GCCCACGTGACGACGAGATCCGCCTCTGCTCAO GGGCAGT	150		
Qy	153	TyrValArgAsnGlyAlaAsnProLeuHis-----GluProValThrGlyHisHis	169		
Db	151	CTCCTCGATGCGGACCGGGGCTCTTCGAGTTGGATCTGAACCATTT-----TACCAC	204		
Qy	170	PhePheAspGlyAspGlyMetValHisalaVallYsPheGluHisGlySerAlaSerTyr	189		
Db	205	CTGTTTTGACGGAACAAGCCCTTCTGCACAA GTTCGACTTTTAAAGAAGGACACGTCACCTAT	264		
Qy	190	AlaCysArgPheThrGlnThrAnrArgPheValGlnGlu-----	202		
Db	265	CACAGAAGGTTTCATCGCACCGATGCTTACGTCGGGCAATGACCGAAAAAGGATCGTC	324		
Qy	203	-- ArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGlyHisThr	221		
Db	325	ATAACGGAAATTGGCACCTGTCGTTCCAGATCCCTGCAAGAATATA-----	372		
Qy	222	GlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaAlaGlyIleValaspProAla	241		
Db	373	---TTTTCCAGGTTTTTTTCTTACTTC-----	396		
Qy	242	HisGlyThrGlyValAlaAsnAlaGlyLeuValTyrPheAsnGlyArgLeuLeuAlaMet	261		
Db	397	CGAGGAGTGGAGTCACTGACAAATGCCCTTGTT-----AACGTCTACCCAGTA	444		
Qy	262	SerGluAsp-----AspLeuProTyrGlnValGlnIleThrProAsnGly	276		
Db	445	GGGGAAGATTACTACGCCTGCACGAGACCAACTTCATTACAAGATTATCTCTGAG--	501		
Qy	277	AspLeuLysThrValGlyArgPheAspGlyGlnLeuGlu---SerThrMetIle	295		
Db	502	ACCCCTGGACAAATTAAGCAGGTGTGATCTCTGCAACTACGTTCTGTCAATGGAGCCACC	561		
Qy	296	AlaHisProLysValAspProcluser-----GlyGlu	306		
Db	562	GCTACCCCCACATTGAANAATGATGGGACTGTTTACAACATTGGTAATTCGTTTGGGAAA	621		
Qy	307	LeuPheAlaLeuSerTyrAspValValSerLysProTyrLeuLysTyrPheArgPheSer	326		
Db	622	AATTTTTTCGATTGCCTACAATATTGTAAAGATCCCTCCACTCCAAGCAGACAGAAGAT	681		
Qy	327	ProAspGlyThrLysSerProaspValGluIleGlnLeuAsp-----Gln	341		
Db	682	CCA-----ATAAGCAAGTCCGAGTCCGTGTACAATTTCCCTTCGACGACCGCATTTCAAG	735		
Qy	342	ProThrMetMetHisAspPheAlaIleThrGluAsnPheValValProAspGlnGln	361		
Db	736	CCATCTAGCTCCATAGTTTGTTTGATTGACTCCCACACTATATTGTTTTTGTGGAGACGCCA	795		
Qy	362	ValValPheLysLeuProGluMetIleArgGlyGlySer-----	374		
Db	796	GTCAAAATTAACCTGCTCAAGTTCTTCTTCG TGAGTCTTTTGGGAGCCCACTACATG	855		
Qy	375	-----ProValValTyrAspLysAsnLysValAla	384		
Db	856	GATTGTTTTGAGTCCAATCAACCATGGGGGTTTGGGCTTCACATCGCTGCAAAAAAAAAA	915		

```
; LENGTH: 1724
; TYPE: DNA
; ORGANISM: Canis familiaris
US-09-645-370-1

Alignment Scores:
  3.3e-13      Length: 1724
Pred. No.:    210.00    Matches: 113
Score:        37.08%    Conservative: 98
Best Similarity: 19.86% Mismatches: 218
Query Match:   6.67%    Indels: 140
DB:            4       Gaps: 25

US-09-758-269-6 (1-599) x US-09-645-370-1 (1-1724)
Qy 120 SerValGlnIle-----AlaGlyAsnPheAlaProValAsnGluInProValArg 136
Db 31 TCCATCCAAAGTGGAGCATCCCGCGGCTTACAAAGAGCTGTTGAAACCGTGAAGAG 90
Qy 137 ArgAsnLeuPro-----ValValGlyLysLeuProAspSerIleLysGlyVal 152
Db 91 CTGTCGTCGCCGCTCACGCCACGTGACAGGAGGATCCCGCTCTCGCTCACGGGCGAGT 150
Qy 153 TyrValArgAsnGlyAlaAsnProLeuHis-----GluProValThrGlyHisHis 169
Db 151 CTCCTCCGATCGGACCGCGGCTCTTCGAGGTTGGATCTGAACCATTT-----TACCAC 204
Qy 170 PhePheAspGlyAspGlyMetValHisAlaValLysPheGluHisGlySerAlaSerTyr 189
Db 205 CTGTTTGACGACAAAGCCCTTCTGCACAAAGTTCGACTTTAAAGAAGGACACGTCACCTAT 264
Qy 190 AlaCysArgPheThrGlnThrAsnArgPheValGlnGlu----- 202
Db 265 CACAGAAGGTTTCATCCCGCACCGATGCTTACGTCGGGCAATGACCGAGAAAGATCGTC 324
Qy 203 ---ArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGlyHisThr 221
Db 325 ATAACGGAAATTGGCACCTGTCGCTCCAGATCCCTCAAGATATA----- 372
Qy 222 GlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaGlyIleValAspProAla 241
Db 373 ---TTTTCCAGGTTTTTTTCTTACTTC----- 396
Qy 242 HisGlyThrGlyValAlaAsnAlaGlyLeuValTyrPheAsnGlyArgLeuLeuAlaMet 261
Db 397 CGAGGAGTGGAGGTCACTGACAATGCCCTTGTT-----AACGTCTACCCAGTA 444
Qy 262 SerGluAsp-----AspLeuProTyrGlnValGlnIleThrProAsnGly 276
Db 445 GGGGAAGATTACTACGCTCGACGAGACCAACTTCATTACAAAGATTAACTCTGAG--- 501
Qy 277 AspLeuLysThrValGlyArgPheAspPheAspGlyGlnLeuGlu---SerThrMetIle 295
Db 502 ACCCTGGAGACAAATTAAAGCAGTTGATCTCTGCAACTACGTCCTGTCAATGGAGCCACC 561
Qy 296 AlaHisProLysValAspProGluSer-----GlyGlu 306
Db 562 GCTCACCCCCACATTGAAATGATGGGAGCTGTTTACAACTTGGTAATGTTTGGGAAA 621
Qy 307 LeuPheAlaLeuSerTyrAspValValSerLysProTyrLeuLysTyrPheArgPheSer 326
Db 622 AATTTCGATTCGCTACAAATATTGTAAGATCCCTCCACTCCAGCAGACAAAGAGAT 681
Qy 327 ProAspGlyThrLysSerProAspValGluIleGlnLeuAsp-----Gln 341
Db 682 CCA-----ATAAGCAAGTCCGAGGTCTCGTACAAATTCCTCCCTGCGAGCGCGATTCAAG 735
Qy 342 ProThrMetMetHisAspPheAlaIleThrGluAsnPheValValProAspGlnGln 361
Db 736 CCATCGTACGTCCTAGTTTGGTTGGTACTCCCACTATATTGTTTGTGGAGACGCCA 795
Qy 362 ValValPheLysLeuProGluMetIleArgGlyGlySer----- 374
```

```
Db 796 GTCAAAATTAACCTGCTCAAGTTCTTTCTTCGTGGAGTCTTTGGGGAGGCAACTACATG 855
Qy 375 -----ProValValTyrAspLysAsnLysValAla 384
Db 856 GATTGTTTTAGTCCCAATGAACATGGGGGTTTGGCTTCACATCGCTGACAAAAGAA 915
Qy 385 ArgPheGlyIleLeuAspLysTyrAlaGluAspSerSerAsnIleLysTyrIleAspAla 404
Db 916 AAAAAGTATCTCAATAATAGTACAGGACCTCTTCCTTTAATCTC----- 960
Qy 405 ProAspCysPheCysPheHisLeuTyrAsnAlaTyrPduGluProGluThrAspGluVal 424
Db 961 -----TTCATCATATCAATACATTACGAGACAATGAGTTT-----CTG 999
Qy 425 ValValIleGlySerCysMetThrProProAspSerIlePheAsnGluSerAsp---Glu 443
Db 1000 ATTGTGGATCTCTGCTGCTGCGAAGGATTTGAATTCGTCTACAATTAATTGTTAGCC 1059
Qy 444 AsnLeuLysSerValLeuSerGluIleArgLeuAsnLeuLysThrGlyGluSerThrArg 463
Db 1060 AATTAGCTGAGAACTGGGAAGAGTGAAA-----AAAAATGCCAGAAAGGCTCCG 1110
Qy 464 ArgProIleIleSerAsnGluAspGlnGlnValAsnLeuGluAlaGlyMetValAsnArg 483
Db 1111 CAGCCTCAAGTTAGGAGATCCGTCTTCTTGAATATCGACAAGGCCGACACAGGCAAG 1170
Qy 484 AsnMetLeuGlyArgLysThrLysPheAlaTyrLeuAlaLeu----- 497
Db 1171 AACCTAGTCACTCCCTCCCAACACGACGCGCACTGCAACTCTGCGACGACGAGACCATC 1230
Qy 498 ---AlaGluProTyrProLysValSerGly-----PheAlaLysVal 510
Db 1231 TGGCTGGAACCTGAGGTTCTTCTCAGGGCCTCGTCAAGCCTTTGAGTTTCTTCAATC 1290
Qy 511 AspLeuThrThr-----GlyGluValLysLysHisLeuTyrGly----- 523
Db 1291 AACTATCAGAAGTATGCGGGAAGCCTTACACGTACGTCGTATGAGACTTGGCTTGAATCAC 1350
Qy 523 ----- 523
Db 1351 TTCGTTCCGGACAGGCTCTGCAAGCTGAACGTCAAGACTAAAGAAACGTTGGTATGCCAA 1410
Qy 524 ---AspAsnArgTyrGlyGlyProLeuLeuLeuPro---GlyGluGlyGlyGluGlu 541
Db 1411 GAGCCCACTCATACCCATCAGAACCCATCTTTGTTTCTCACCAGATGCTTGGAGAA 1470
Qy 542 AspGluGlyTyrIleLeuCysPheValHisAspGluLysThrTrpLys-----SerGlu 559
Db 1471 GATGATGGTGTAGTTCTGAGTGTGTTGAGCCCTGGGCGAGGACAAAGACCTGCTTAT 1530
Qy 560 LeuGlnIleValAsnAlaValSerLeuGluValGluAlaThrValLysLeuProSerArg 579
Db 1531 CTTCTGATTCTGAATGCCAAGGATTTGAGTGAAGTTCCAGGGCTGAAGTGGAGATTAAAC 1590
Qy 580 ValProTyrGlyPheHisGlyThrPhe 588
Db 1591 ATCCCTGTCACCTTTCATGACTGTTT 1617

RESULT 11
US-08-311-731A-140
; Sequence 140, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
```

STATE: MASSACHUSETTS  
 COUNTRY: USA  
 ZIP: 02210  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC Compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/311,731A  
 FILING DATE:  
 CLASSIFICATION: 530  
 ATTORNEY/AGENT INFORMATION:  
 NAME: GATES, EDWARD R.  
 REGISTRATION NUMBER: 31,616  
 REFERENCE/DOCKET NUMBER: C0044/7125  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617/720-3500  
 TELEFAX: 617/720-2441  
 INFORMATION FOR SEQ ID NO: 140:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 36063 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: circular  
 MOLECULE TYPE: DNA (genomic)  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Mycobacterium leprae  
 US-08-311-731A-140

Alignment Scores:  
 Pred. No.: 2,74e-08 Length: 36063  
 Score: 186.00 Matches: 133  
 Percent Similarity: 38.61% Conservative: 72  
 Best Local Similarity: 25.05% Mismatches: 209  
 Query Match: 5.90% Indels: 121  
 DB: 4 Gaps: 23

US-09-758-269-6 (1-599) x US-08-311-731A-140 (1-36063)

QY 110 LeuHisProLeuProLysThrAlaAspProSerValGlnIleAlaGlyAsnPheAlaPro 129  
 DB 22576 TTGTGACGACTATAGAGGATCAGCACATGGATGTGCGAATTTGTGAGCAAGTCTGTCT 22635  
 QY 130 Val-----AsnGluClnProValArgArgAsnLeuPro----- 140  
 DB 22636 ACCCTGTCCGAGACGACGACCTCGCGCGGATTTGTCCTGCGCGACCGACGATC 22695  
 QY 141 -----ValValGlyLysLeuProAspSerIleLysGly 151  
 DB 22696 GNAATGGGAGCCACGACCTCAACGCTTTAGCGGGAGAAATACCCAGTACTCTAAAGCG 22755  
 QY 152 ValTyrValArgAsnGlyAlaAsnProLeuHisGluProValThrGlyHisPhePhe 171  
 DB 22756 TACCTACCGACGCAACACTAAACCCACTGCAACCGGATTCACAGTGTATCACCCGTT 22815  
 QY 172 AspGlyAspGlyMetValHisAlaValLysPheGluHisGlySerAlaSerTyrAlaCys 191  
 DB 22816 GATGAAGACGGCATGCTGCACATGGTCGCTTCCGCGATGGAAAGCCCTTTTTCGCAAC 22875  
 QY 192 ArgPheThrGlnThrAsnArgPheValGlnGluArgGlnLeuGlyArgProValPhePro 211  
 DB 22876 AGCTTTATACGACAGACGATTTTGTCCGAGAACATCGAGCGCGCTGTGGCC 22935  
 QY 212 LysAlaIleGlyGluLeuHisGlyHisThrGlyIleAlaArgLeuMetLeuPheTyrAla 231  
 DB 22936 GG-TTG-----GGCAGAACCGGT----- 22952  
 QY 232 ArgAlaAlaAlaGlyIleValAspProAla---HisGlyThrGlyValAlaAsnAlaGly 250  
 DB 22953 -----GCAATAATTGACCAAGCGTGAACAGGCGCTGGGGATCCGCAACGATG 23000

QY 251 LeuValTyrPheAsn-----GlyArgLeuLeuAlaMetSerGlu----- 263  
 DB 23001 AAGGACGGTTGCAACACCGACGCTACCGCTCCATCGAGCATCGCGCTGACCATTTCTTAC 23060  
 QY 264 -----AspAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysThrVal 281  
 DB 23061 CATTGACGCGATCTG---TACCACATAGAACCCTATCCG-----GCCAATACGCGA 23108  
 QY 282 GlyArgPheAspPheAspGlyGlnLeuGluSerThr---MetIle-AlaHisProLysVal 300  
 DB 23109 GCGAAGGAGATCGGGGCGCGCTTACAGTTGACTGGAGCGTGGCGGCACATCTCAAACT 23168  
 QY 300 LaspProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValSer-LysProTyrL 320  
 DB 23169 GGACAAACGAGACCGGAGAACTG-----CTGTTCTTTAATACAAAGCAGGACCGGTACA 23222  
 QY 320 euLysTyrPheArgPheSerProaspGlyThrLys---SerProaspValGluIleGlnL 339  
 DB 23223 CATCGCTATGCGCTCAACCAACAAAGCAAGCACTGACACACACACAGATCGCATTTCCG 23282  
 QY 339 euAspGlnProThrMetMetHisAspPheAlaIleThrGluAsnPheValValProA 359  
 DB 23283 TGCTTGGTTCGGTTTGCCACATGGCGTTCACCGAAACACTACAGATCCTCAATA 23342  
 QY 359 spGlnGlnValValPheLysLeuProGluMetIleArgGlyLysSerProValValTyrA 379  
 DB 23343 ATTG-CCCATTTGTTCTGAAG-----TCCAAGACTGCTCAAG 23377  
 QY 379 spLysAsnLysValAlaAlaArgPhe-----GlyIleLeuAspLysTyrAlaGluAspSers 397  
 DB 23378 AACAAATGTGTCTACCCAGATTTCTATCTGTCATGTCGTCTCTCGTTCGCGTGT-TCAA 23436  
 QY 397 er-----AsnIleLysTrpIleAspAlaProAspCysPheCysPheHisLeuTrpAsnA 415  
 DB 23437 GCACCGGCGACATCCGGTGATTCGAGGCGGAACTTATATCTTACCGCATTTTCTCAAG 23496  
 QY 415 latrGluGluProGluThrAspGluValValIleGlySerCysMetThrProProA 435  
 DB 23497 CCTACGAA-----AAGGGTCACGAGATCATG-----CGCG 23526  
 QY 435 spSerIlePheAsnGluSerAspGluAsnLeu-----LysSerValL 449  
 DB 23527 ATAGATTTCTCGAAGACTAGCCACACCCCTTTGACGCGCGGCCACTGACCAAGCCCGG-C 23585  
 QY 449 euSerGluIleArgLeuAsnLeuLysThrGlyGluSerThrArgArgProIleIleSerA 469  
 DB 23586 TGCAAGCTGCGGATTTCACTACCGGTGCTGTAAACAGAGAGTAACCTG----- 23638  
 QY 469 snGluAspGlnGlnValAsnLeuGluAlaGlyMetValAsnArgAsnMetLeuGlyArgL 489  
 DB 23639 -----CCGGAATCTATCACCGAGTTCGGAGTATCAATGCTGACTATGCCTCCAGCA 23690  
 QY 489 ysThrLysPheAlaTyrLeuAlaLeuAlaGluProTrpProLysValSerGly----- 506  
 DB 23691 AGTATCACTACTCTATACCTATGCGGC-----ACCGGCCAATCGGGCTAGTTCT 23741  
 QY 507 -----PheAlaLysValAspLeuThrThrGlyGluVal-LysLysHisLeuTyr 522  
 DB 23742 TCTTAGCGTATGCTTAACACCGTTCTGTCGACCAACACCTAGGAGCGCTACTCATTC 23801  
 QY 523 GlyAspAsnArgTyrGlyGlyProLeuPheLeuProGlyGlyGlyGlyGlu---Glu 541  
 DB 23802 GCGCATGCATCTACGGAAGTGAGAGGCGGTGCTCGCGGTGGGTAGTACCGGTGAA 23861  
 QY 542 AspGluGlyTyrIleLeuCysPheValHisAspGluLysThrTrpLys----- 557  
 DB 23862 GATACGCGCTAACTGGTAAACCCCTACCAACCGACATGAATGCCGATAAGTCTTAATATAT 23921  
 QY 558 -----SerGluLeuGlnIleValAsn 564  
 DB 23922 GGGCTTCGAAGCAGCCAGGATCACCGACACGTCGAGGTGAGTAAACTCAACTGCGGAAA 23981

Qy 565 AlaValSerLeuGluValGluAlaThr 573  
Db 23982 CGTATTCACGAGCGCGCTTCCACC 24008

RESULT 12  
US-09-252-991A-7056/c  
; Sequence 7056, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 7056  
; LENGTH: 4242  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-7056

Alignment Scores:  
Pred. No.: 0.00236 Length: 4242  
Score: 127.50 Matches: 104  
Percent Similarity: 35.96% Conservative: 65  
Best Local Similarity: 22.13% Mismatches: 174  
Query Match: 4.05% Indels: 129  
DB: 4 Gaps: 21

US-09-758-269-6 (1-599) x US-09-252-991A-7056 (1-4242)

Qy 23 ProLeuSerSerGlnSerSerAspLeuSerTyrCysSerSerLeuProMetAlaSer 42  
Db 2486 CCATGCACTGCGCGGTGATCGAGTCTTGGCATGC-----GCGCCG 2442

Qy 43 ArgValThrArgLysLeuAsnValSerSerAlaLeuHisThrProProAlaLeuHisPhe 62  
Db 2441 AGGATTGGAACTGCACCTTCTATTCTGATCACTTCCAGCGCTGCCAGCGCGCTGTGG 2382

Qy 63 ProLysGlnSerSerAsnSerProAlaLeuValLysProLysAlaLysGluSerAsn 82  
Db 2381 CCCCCTTGTCTGTGC--GCGCCCGCGTGTGTCGCGGCGCCAGGCGCAGTGGGCGCGG 2325

Qy 83 ThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaLeuAspAlaAlaGluGly 102  
Db 2324 AGGAAA-----TCTGGGAG-CTGATCCGCGCGGAGGCG 2293

Qy 103 PheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSerValGln 122  
Db 2292 -GTGAG-CATCTCTCGCTTCCACCCGAGCTACGG----- 2261

Qy 123 IleAlaGlyAsnPheAlaProValAsnGluGlnProValArgArgAsnLeuProValVal 142  
Db 2260 -----CAGCCAGCTCGC----- 2240

Qy 143 GlyLysLeuProAspSerIleLysGlyValTyrValArgAsn-GlyAlaAsnProLeuHi 162  
Db 2239 GGAGAGCCAGCGCGGATGTCGGTGGCGATGTCATCATCCGCGGC----- 2191

Qy 162 sGluProValThrGlyHisHis-----ph 170  
Db 2190 -GAAGCGCTGACCGCGAGACCTGCAACGAGTTCGCCAGGCCTTGGCCCGCGGTCTT 2132

Qy 170 ePheAspGlyAspGlyMetValHisAlaVal----- 180  
Db 2131 CTTCAATGCCCTACGACCGACCGAGAGCGTGGTGCATGCCGTGCGCTGCTGCCGA 2072

Qy 181 -LysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheVa 200

Db 2071 ACGACTCGAGAGGGTCCGCCACGCGTCCGATCGGCAGCGTGGTTCGCGCGCGGTCGC 2012  
Qy 200 lGlnLysArgGlnLeuGlyArgProValPheProLys--AlaIleGlyGluLeuHis-- 218  
Db 2011 CTACATCTCGATGCCACCTGGCCCTTGGTGGTGGCGGCGCGGCGGCGGCGGCGGCGG 1952  
Qy 219 -GlyHisThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaGly----- 236  
Db 1951 CGGCGCGCGCGCGCTGGCGGC-----GGCTACCATGAGCGTCCGCGCTCAGCGCGGA 1898  
Qy 237 -----IleValAspProAlaHisGlyThrGlyValAlaAsnAlaGlyLeuValTyrPh 254  
Db 1897 GCGCTTCGTGCCGATCCCTTC-----GCTGCCGAGGCGCGCGCTGATCCG 1850  
Qy 254 eAsnGlyArgLeuLeuAlaMetSerGluAspAspLeuProTyrGlnValGlnIleThrPr 274  
Db 1849 CACGCGGACCTG-----GTGCGCTGTGCGA 1823  
Qy 274 oAsnGlyAspLeuLysThrValGlyArgPheAspPheAsp----- 287  
Db 1822 CACGCGGAGTGAATATGTCCGCCCATCGACCCAGGTGAAGATCCGTGCTTCCG 1763  
Qy 288 -----GlyGlnLeuGluSerThrMetIleAlaHisProLysValAspProGluSe 304  
Db 1762 CATCGAACTGGCGAGATCGAGCGCGCTGTGGAGCATCCGCGAGTT----- 1714  
Qy 304 rGlyGluLeuPheAlaLeuSerTyrAspValValSerLysPheProTyrLeuLysTyrPheAr 324  
Db 1713 -CGCAAGCGCTGTCTGCTGCGTGCAGACCCGCGGCAAGCAGTTGGCGGCTTATGT 1655  
Qy 324 gPheSerProAspGlyThrLysSerProAspValGluLeuGlnLeuAspGlnProThrMe 344  
Db 1654 CGCAGCGCGGTGCGCGAGCAGGACGCGCGCTGCGCGCGCTGCGGAGCGCGCTGAA 1595  
Qy 344 tMetHisAspPheAlaIleThrGluAsnPheValValProAspGlnGlnValPh 364  
Db 1594 AACGCATCTCAAGCAGCAATTGCGGACTACATGCTGCGCGCCACCTGCTGTGCTCGC 1535  
Qy 364 eLysLeuProGluMetIleArgGlyGlySerProValValTyrAspLysAsnLysValAl 384  
Db 1534 CAGCTTCCCG-----CTGAC 1520  
Qy 384 aArgPheGlyIleLeuAspLysTyrAla-----GluAspSerSerAsnIleLy 400  
Db 1519 CGCCAAACGGAAGCTCAGCCGCGCGCTGCGCGCGCGCGCGCGCGCGCTCAACCGGCA 1450  
Qy 400 sTrpIleAspAlaPro---AspCysPheCysPheHisLeuTrpAsnAlaTrpGluPr 419  
Db 1459 GGCCTACGAGCGCGCGCGAGCGTGGAGCAGCAACTGCGCGGGGTCTGGCGGAGGT 1400  
Qy 419 oGluThrAspGluValValIleGlySerCysMetThrProAspSerIlePheAs 439  
Db 1399 GCTGAACGTCGCGGGTAGGTCTCGGC-----GACAACTTCTTCGA 1358  
Qy 439 nGluSerAspGluAsnLeuLysSerVal 448  
Db 1357 ACTGGCGCGCATTCGATCTGTGATC 1330

RESULT 13  
US-09-252-991A-6997  
; Sequence 6997, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190



;; PRIOR FILING DATE: 1998-07-27

;; NUMBER OF SEQ ID NOS: 33142

;; SEQ ID NO 6997

;; LENGTH: 10023

;; TYPE: DNA

;; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-6997

# Alignment Scores:

Pred. No.: 0.0102 Length: 10023  
Score: 127.50 Matches: 104  
Percent Similarity: 35.96% Conservative: 65  
Best Local Similarity: 22.13% Mismatches: 174  
Query Match: 4.05% Indels: 129  
DB: 4 Gaps: 21

US-09-758-269-6 (1-599) x US-09-252-991A-6997 (1-10023)

QY 23 ProLeuSerSerSerClnSerSerAspLeuSerTyrCysSerSerLeuProMetAlaSer 42  
Db 6671 CCATGCACCTGCGCGCGGTGATCGAGTGTCTCGCATGC-----GCGCG 6715  
QY 43 ArgValThrArgLysLeuAsnValSerSerAlaLeuHisThrProProAlaLeuHisPhe 62  
Db 6716 AGGATTGCGAATGCACTTCTATTTCATCACTTCGACCTGCCAGCGAGCGCTGCTGG 6775  
QY 63 ProLysGlnSerSerAsnSerProAlaIleValLysProLysAlaLysGluSerAsn 82  
Db 6776 CCGCGTGTGTGCG---CGCGCGCGTGTGTGCGCGCGCGAGGCCAGTGGCGCGG 6832  
QY 83 ThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaLeuAspAlaGluGly 102  
Db 6833 AGGAAA-----TCTCGAG-CTGATCGCGCGAGGCG 6864  
QY 103 PheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSerValGln 122  
Db 6865 -----GTGAG-CATCTCGGTTTACCCCGAGCTACGG----- 6896  
QY 123 IleAlaGlyAsnPheAlaProValAsnGluGlnProValArgAsnLeuProValVal 142  
Db 6897 -----CAGCCAGCTCGC-----CCAGTGGCT 6917  
QY 143 GlyLysLeuProAspSerIleLysGlyValTyrValArgAsn-GlyAlaAsnProLeuHi 162  
Db 6918 GGAGAGCCAGGCGCGAGTTCGCGTGCATGTGTCATCACCGCGGC----- 6966  
QY 162 sglProValThrGlyHisHis-----Ph 170  
Db 6967 -GAAGCGTGCACCGCGAGCACCTGCAACGGATTGCCAGGCTTCGCCCGCGCGTGT 7025  
QY 170 ePheAspGlyAspGlyMetValHisAlaVal----- 180  
Db 7026 CTTCAATGCTACGAGACCGAGAGCGGTGTGTCATCGCGTGGCTGCTGCTGCCCGA 7085  
QY 181 -LysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheVa 200  
Db 7086 ACAGCTGGAGGAGGTGCGCGCAGCGTGCATCGCGAGCGTGTGCGCGCGCGGTGCG 7145  
QY 200 lGlnGluArgGlnLeuGlyArgProValPheProLys---AlaIleGlyGluLeuHis-- 218  
Db 7146 CTACATCTCGATGCGGACCTGGCGTGTGCGGAGGCGCGCGGCGGAACTCTACGT 7205  
QY 219 -GlyHisThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaAlaGly----- 236  
Db 7206 CGCGCGCGCGCGCTGCGCGC-----GGCTACCATGAGCTGCGCGGCTCAGCGCGCA 7259  
QY 237 -----IleValAspProAlaHisGlyThrGlyValAlaAsnAlaGlyLeuValTyrPh 254  
Db 7260 GCGCTTGTGCGCGATCCCTTC-----GCTGCCAGGCGCGCGCTGTACCG 7307  
QY 254 eAsnGlyArgLeuLeuAlaMetSerGluAspAspLeuProTyrGlnValGlnIleThrPr 274  
Db 7308 CACCGCGGACCTG-----GTGCGCGCTGTGCGA 7334

QY 274 oAsnGlyAspLeuLysThrValGlyArgPheAspPheAsp----- 287  
Db 7335 CAACGCGCAGGTGAATATGTGCGCGCATCGACCAAGATCGATCGGTGCTTCG 7394  
QY 288 -----GlyGlnLeuGluSerThrMetIleAlaHisProLysValAspProGluSe 304  
Db 7395 CATCGAACTGGCGGAGATCGAGCGCGCTGTGTGAGCATCGCAGGTT----- 7443  
QY 304 rGlyGluLeuPheAlaLeuSerTyrAspValValSerLysProTyrLeuLysTyrPheAr 324  
Db 7444 -CGCGAAGCGCTGCTGCTGCGCTCGACAGCCCGAGCGCAGCAGTGGCCGGTATGT 7502  
QY 324 gPheSerProAspGlyThrLysSerProAspValGluIleGlnLeuAspGlnProThrMe 344  
Db 7503 CGCCAGCGCGTGGCGGAGCAGGAGCGCCAGCGCGCGCTGCGCGAGCGGTGAA 7562  
QY 344 tMethHisaspPheAlaIleThrGluAsnPheValValProAspGlnValValPh 364  
Db 7563 AACGCATCTCAAGCAGCAATTGCGGACTACATGTGCGGCGCCACCTGCTGCTCGC 7622  
QY 364 eLysLeuProGluMetIleArgGlySerProValValTyrAspLysAsnLysValAl 384  
Db 7623 CAGCTGCGC-----CTGAC 7637  
QY 384 aArgPheGlyIleLeuAspLysTyrAla-----GluAspSerSerAsnIlely 400  
Db 7638 CGCCAAAGCGAAGCTCGACCGCGCGCTGCGCGCGCGCGCGCTCAACCGGCA 7697  
QY 400 sTrpIleAspAlaPro---AspCysPheCysPheHisLeuTrpAsnAlaTrpGluGluPr 419  
Db 7698 GCGCTACAGCGCGCGCGCGCGCTGCTCGAGCAGCAACTGGCGCGGTCTGCGCGAGGT 7757  
QY 419 oGluThrAspGluValValIleGlySerCysMetThrProProAspSerIlePheAs 439  
Db 7758 GCTGACCTCGAGCGGTAGGTCTCGGC-----GACAACTTCTTGA 7799  
QY 439 nGluSerAspGluAsnLeuLysSerVal 448  
Db 7800 ACTGGCGCGATTTCGATCTGTCGATC 7827

## RESULT 14

US-08-700-651-1  
; Sequence 1, Application US/08700651B  
; Patent No. 6015882

## GENERAL INFORMATION:

; APPLICANT: PETERSEN, CAROLYN  
; APPLICANT: LEECH, JAMES  
; APPLICANT: NELSON, RICHARD, C.  
; APPLICANT: GUT, JIRI  
; TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS  
; TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum  
; TITLE OF INVENTION: INFECTIONS  
; FILE REFERENCE: 480.19-4 (HV)  
; CURRENT APPLICATION NUMBER: US/08/700.651B  
; CURRENT FILING DATE: 1997-08-14  
; EARLIER APPLICATION NUMBER: 08/415.751  
; EARLIER FILING DATE: 1995-04-03  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 5163  
; TYPE: DNA  
; ORGANISM: Cryptosporidium parvum  
US-08-700-651-1

## Alignment Scores:

Pred. No.: 0.0092 Length: 5163  
Score: 123.50 Matches: 125  
Percent Similarity: 34.60% Conservative: 85  
Best Local Similarity: 20.59% Mismatches: 194  
Query Match: 3.92% Indels: 203  
DB: 3 Gaps: 35

US-09-758-269-6 (1-599) x US-08-700-651-1 (1-5163)

QY 5 ThrAlaThrAlaAlaValSerGlyArgTrpLeuGlyGlyAsnHisThrGlnProLeu 24  
DB 2491 ACAATTGCAGGTATTGTTTCAGGA-----ATT 2517

QY 25 SerSerGlnSerSerAspLeuSerTyrCysSerSerLeu-----ProMetAlaSer 42  
DB 2518 TCTGCAAGTGAGTCAATTA---TTATCTCAGAAATCAGCTCTAATCGGCCAGCAACAAAT 2574

QY 43 ArgValThrArgLys-----LeuAsnValSerSerAlaLeuHisThrProPro 58  
DB 2575 ATGTTGTTGGAGAAATTTGGTGGATTGTTGAACCCAGCAACAGGAGTGATCCAGGT 2634

QY 59 AlaLeuHisPheProLysGlnSerSerAsnSerProAlaIle-----ValVal 74  
DB 2635 TTTTGTAGTCCATCAGAGCAAACTCAATCTCCCTGAGATTGAAGATTGGTGTATTATT 2694

QY 75 LysProLysAlaLysGluSerAsnThrLysGlnMetAsnLeu-----88  
DB 2695 CCTCCAGAAAGTAGCAGCAAACTGCTGATAAATTCAGATTATCTATTCTCCCAAGCGTA 2754

QY 89 -----PheGlnArgAlaAlaAlaAlaLeuAsp 98  
DB 2755 CCAGAAATCAATTCAGAAAAGGATCAGAAAGATTGATTCTATTCTGAATTGATGATGAT 2814

QY 99 AlaAlaGluGlyPheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAsp 118  
DB 2815 ATTGAGTCAGGTAGACTTATTGGTCAAGTATCAAGAGACCAATCCAGGTTC-----2868

QY 119 ProSerValGlnIleAlaGlyAsnPheAlaProValAsnGluGlnProValArgArgAsn 138  
DB 2869 -----ATTGCTGGTGACTTGAACCCCAATAATGAAGACACCAACAACACTGAC 2916

QY 139 LeuProValValGlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyAla 158  
DB 2917 ---AGTGAATGTTGAACCAATCGATCCAACCCAGGTCTG-----2955

QY 159 AsnProLeuHisGluProValThrGlyHisHisPhePheAspGlyAspGlyMetValHis 178  
DB 2956 ---CCTTTCAATCCACCA---ACTGGTCAT-----TTGATTAAC 2988

QY 179 AlaValLysPheGluHisGlySerAlaSerTyr-----AlaCysArgPheThrGlnThr 196  
DB 2989 CCAACAAATAATAATACCATGGATTCCTTCATTGCTGTGTCATCAAAATATGCGATTCA 3048

QY 197 AsnArgPheValGlnGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGlu 216  
DB 3049 AATGGTATTAGACTGATAATGTTTATGTTTACCAGTT-----GGTGAA 3093

QY 217 LeuHisGlyHisThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaGly 236  
DB 3094 ATAACAGGTTTACCAGGATCCAGGCTCAGATATTCCATTAACTCAACTACAGGTGAA 3153

QY 237 IleValAspProAlaHisGlyThrGlyValAlaAsn-----AlaGlyLeuValTyrPhe 254  
DB 3154 TTAGTTGATCCATCAACAGGAAGCAATTAACAAATCTATGCTGGTATTGTT-----3207

QY 255 AsnGlyArgLeu---LeuAlaMetSerGluAspAspLeuProTyrGlnValGlnIleThr 273  
DB 3208 AGTGAAAACCTGGCTTACCACCTATTGAAGTGA-----3243

QY 274 ProAsnGlyAspLeuLysThrValGlyArgPheAsp-----PheAsp 287  
DB 3244 ---AATGTAATTG-----TTTGATCCATCAACTCAACTGCAATAGAT 3285

QY 288 GlyGlnLeuGluSerThrMetIleAlaHisProLysValAspProGluSerGlyGluLeu 307  
DB 3286 GGTAAAT-----AACCAATTAGTTAAACCAAGAACCAACAGCACT 3324

QY 308 PheAlaLeuSerTyrAspValValSerLysProTyrLeuLysTyrPheArgPheSerPro 327  
DB 3324 -----

DB 3325 GTCTCAGGATCAACTTCAGGTACTACAAAACCA-----AAACCA 3363

QY 328 -----AspGlyThrLysSerProAspValGluIle-----GlnLeuAsp 340  
DB 3364 GGAATTCAGTCAATGTGGAGGTGTGTACTCTGATGAAGAAGCTAAAGATCAAGCCGAT 3423

QY 341 Gln-----ProThrMet 344  
DB 3424 AAGGTAAGGATGATTAAATTTGTTCCCAACTAATTTCTATCAATAAAGATCCAGTAACA 3483

QY 345 MethHisAspPheAlaIleThrGluAsnPheValValProAspGlnGlnValPhe 364  
DB 3484 AATACTCAGTACAGTAATACTACTGTTAAACATTATTAAACCCAGAAACAGGAAAAAGTT---3540

QY 365 LysLeuProGluMetIleArgGly-----GlySerProValVal 377  
DB 3541 ---ATTCCAGGTTCCACTCCAGGCTCTCTCAACTATCCATCAATTAATCTCCCAACAA 3597

QY 378 TyrAspLysAsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSerSer 397  
DB 3598 ACTGATGAG-----ATTACAGAAAGCCAGTTGATCTGTTACT 3636

QY 398 AsnIleLysTrpIleAspAlaProAspCysPheCysPheHisLeuTrpAsnAlaTrpGlu 417  
DB 3637 GGTTCGCATAT-----3648

QY 418 GluProGluThrAspGluValValIleGlySerCysMetThrProProAspSerIle 437  
DB 3649 GATCCATCTCAGGTGAAATTTATCGATCCTGCAACTAAATTTACCAATTTCCAGGATCAGTT 3708

QY 438 PheAsnGluSerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsnLeuLys 457  
DB 3709 GCAGGTGAT-----GAAATCTCTACTGAAGTA---TTGAACATTACA 3747

QY 458 ThrGlyGluSerThrArgArgProIleIleSerAsnGluAspGlnGlnValAsnLeuGlu 477  
DB 3748 ACAGATGAAGTAACAGGTTTGCCCAATT-----GATCTTGAA 3783

QY 478 AlaGlyMetValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeuAlaLeu 497  
DB 3784 ACTGGTCTT-----3792

QY 498 AlaGluProTrpProLysValSerGlyPheAlaLysVal-----AspLeuThrThr 514  
DB 3793 -----CCAAGATCCAGTATCAGGATCCCAACTCCCAACTTCCAAATGGTTACTTGGTTGAT 3846

QY 515 GlyGluValLysLysHisLeuTyrGlyAspAsnArgTyrGlyGlyGluProLeuPheLeu 534  
DB 3847 CCATCAAAATAAAACCAATTCAGGTTC-----CATTCGGA-----TTTATT 3891

QY 535 ProGlyGluGlyGlyGluGlu 541  
DB 3892 AATGGTATCATCTGGGAACAA 3912

RESULT 15  
US-08-928-361B-4  
; Sequence 4, Application US/08928361B  
; Patent No. 6071518  
; GENERAL INFORMATION:  
; APPLICANT: Petersen, Carolyn  
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,  
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS  
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM  
; TITLE OF INVENTION: SPECIES INFECTIONS  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: PETERS, VERNY, JONES & BIKSA  
; STREET: 365 Sherman Avenue, Suite 6  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306-1840  
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/928,361B  
 ; FILING DATE: 12-SEP-1997  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 60/026,062  
 ; FILING DATE: 13-SEP-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Vetry, Hana  
 ; REGISTRATION NUMBER: 30,518  
 ; REFERENCE/DOCKET NUMBER: 480.76-1 (HV)  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 650-324-1677  
 ; TELEFAX: 650-324-1678  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 5163 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; US-08-928-361B-4

Alignment Scores:  
 Pred. No.: 0.0092 Length: 5163  
 Score: 123.50 Matches: 125  
 Percent Similarity: 34.60% Conservative: 85  
 Best Local Similarity: 20.59% Mismatches: 194  
 Query Match: 3.92% Indels: 203  
 DB: 3 Gaps: 35

US-09-758-269-6 (1-599) x US-08-928-361B-4 (1-5163)

QY	159	AsnProLeuHisGluProValThrGlyHisHisPheAspGlyMetValHis	178
DB	2956	---CCTTTCAATCCACCA---ACTGGTCAT	2988
QY	179	AlaValLysPheGluHisGlySerAlaSerTyr	196
DB	2989	CCAACAATAATAATACCATGATTCTTCATTGCTGGTGCATACAAATATGACAGTTTCA	3048
QY	197	AsnArgPheValGlnGlnLeuGlyArgProValPheProLysAlaIleGlyGlu	216
DB	3049	AATGGTATTAAAGACTGATAATCTTATGGTTTACCAGTT	3093
QY	217	LeuHisGlyHisThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaGly	236
DB	3094	ATAACAGGTTTACCAAGGATCCAGGCTCAGATTTCCATTTAACTCAACTACAGGTGAA	3153
QY	237	IleValAspProAlaHisGlyThrGlyValAlaAsn	254
DB	3154	TTAGTTGATCCCATCAACGGAAGCAATTAACAATTTCTACTGCTGGTATTGTT	3207
QY	255	AsnGlyArgLeu---LeuAlaMetSerGluAspAspLeuProTyrGlnValGlnIleThr	273
DB	3208	ACTGGAAACCTGGCTTACCACCTATTGAAGATGAA	3243
QY	274	ProAsnGlyAspLeuLysThrValGlyArgPheAsp	287
DB	3244	---AATGGTAATTG---TTTGATCCATCACTAACTTGCCTAATAGAT	3285
QY	288	GlyGlnLeuGluSerThrMetIleAlaHisProLysValAspProGluSerGlyGluLeu	307
DB	3286	GGTAAT---AACCAATTAGTTAACCCAGAAACCAACAGCAGCT	3324
QY	308	PheAlaLeuSerTyrAspValValSerLysProTyrLeuLysTyrPheArgPheSerPro	327
DB	3325	GTCTCAGGATCACTTCAGGTACTACAAAACCA	3363
QY	328	---AspGlyThrLysSerProAspValGluIle	340
DB	3364	GAATTTCCAGTCAATGGTGGAGGTGTGTACCTGATGAAGAGTAAAGATCAAGCCGAT	3423
QY	341	Gln	344
DB	3424	AAGGTAAAGGTAGGTAATTAATGTTCCACCACTAATTTCTATCAATAAAGATCCAGTAACA	3483
QY	345	MetHisAspPheAlaIleThrGluAsnPheValValProAspGlnValPhe	364
DB	3484	ATACTCAGTACAGTAATACTACTCGTAACATTATTAACCCAGAAACAGGAAGTT	3540
QY	365	LysLeuProGluMetIleArgGly	377
DB	3541	---ATTCCAGGTTCACTCCAGGCTCTCTCAACTATCCATCATTTCAATATCTCCACAA	3597
QY	378	TyrAspLysAsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSerSer	397
DB	3598	ACTGATGAG---ATTACAGGAAAGCCAGTTGATGATCTGTTACT	3636
QY	398	AsnIleLysTyrIleAspAlaProAspCysPheHisLeuTyrAsnAlaTrpGlu	417
DB	3637	GGTTGGCCATAT	3648
QY	418	GluProGluThrAspGluValValIleGlySerCysMetThrProAspSerIle	437
DB	3649	GATCCATCTACAGGTGAAATTTATCGATCTCGCACTAAATTTACCAATTTCCAGATCAGTT	3708
QY	438	PheAsnGluSerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsnLeuLys	457
DB	3709	CCAGGTGAT---GAAATCTCTCACTGAAGTA---TTGAACATTACA	3747
QY	458	ThrGlyGluSerThrArgArgProIleIleSerAsnGluAspGlnValAsnLeuGlu	477
DB	3748	ACAGATGAAGTAAACAGGTTTCCCAATT	3783
QY	478	AlaGlyMetValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeuAlaLeu	497

```

Db      3784 ACTGCTCTT----- 3792
QY      498 AlaGluProTrpProLysValSerGlyPheAlaLysVal-----AspLeuThrThr 514
Db      3793 -----CCAAGAGATCCAGTATCAGGACTCCCAAACTTCCAAATGGGTACCTTGGTTGAT 3846
QY      515 GlyGluValLysLysHisLeuTyrGlyAspAsnArgTyrGlyGlyGluProLeuPheLeu 534
Db      3847 CCATCAAAATAAAAAACCAATTCCAGGTTCA-----CATTCCGGA-----TTTATT 3891
QY      535 ProGlyGluGlyGlyGluGlu 541
Db      3892 AATGGTACATCTGGAGAACAA 3912

```

Search completed: November 15, 2003, 20:02:48  
 Job time : 14378 secs

GenCore version 5.1.6  
- Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 15, 2003, 15:07:29 ; Search time 535 Seconds  
(without alignments)  
3659.167 Million cell updates/sec

Title: US-09-758-269-6  
Perfect score: 3150  
Sequence: 1 MASFTATAAVSCRWLGNNHT.....VPYGFHGTFIGADDLAKQVV 599

Scoring table:  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2169961 seqs, 1634102185 residues

Total number of hits satisfying chosen parameters: 4339922

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO\_spool/US09758269/runat\_14112003\_135800\_28838/app.query.fasta\_1.775  
-DB=Published Applications NA -QFWT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MAIRX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09758269 @CIGN 1.1 347 @runat\_14112003\_135800\_28838  
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2\_6/ptodata/1/pubna/US07\_PUBCOMB.seq:
- 2: /cgn2\_6/ptodata/1/pubna/PCT\_NEW\_PUB.seq:
- 3: /cgn2\_6/ptodata/1/pubna/US05\_NEW\_PUB.seq:
- 4: /cgn2\_6/ptodata/1/pubna/US06\_PUBCOMB.seq:
- 5: /cgn2\_6/ptodata/1/pubna/US07\_NEW\_PUB.seq:
- 6: /cgn2\_6/ptodata/1/pubna/US08\_PUBCOMB.seq:
- 7: /cgn2\_6/ptodata/1/pubna/US08\_NEW\_PUB.seq:
- 8: /cgn2\_6/ptodata/1/pubna/US08\_PUBCOMB.seq:
- 9: /cgn2\_6/ptodata/1/pubna/US09A\_PUBCOMB.seq:
- 10: /cgn2\_6/ptodata/1/pubna/US09B\_PUBCOMB.seq:
- 11: /cgn2\_6/ptodata/1/pubna/US09C\_PUBCOMB.seq:
- 12: /cgn2\_6/ptodata/1/pubna/US09\_NEW\_PUB.seq:
- 13: /cgn2\_6/ptodata/1/pubna/US10A\_PUBCOMB.seq:
- 14: /cgn2\_6/ptodata/1/pubna/US10B\_PUBCOMB.seq:
- 15: /cgn2\_6/ptodata/1/pubna/US10\_NEW\_PUB.seq:
- 16: /cgn2\_6/ptodata/1/pubna/US60\_NEW\_PUB.seq:
- 17: /cgn2\_6/ptodata/1/pubna/US60\_PUBCOMB.seq:

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	Score	Match	Length	ID	Description
1	3150	100.0	1800	10	US-09-758-269-5	Sequence 5, Appli

2	2280.5	72.4	1818	10	US-09-758-269-15	Sequence 15, Appli
3	2168.5	68.8	1839	10	US-09-758-269-11	Sequence 11, Appli
4	1991	63.2	1752	10	US-09-758-269-1	Sequence 1, Appli
5	1930	61.3	1815	10	US-09-758-269-13	Sequence 13, Appli
6	1663.5	52.8	1734	10	US-09-758-269-9	Sequence 9, Appli
7	969	30.8	1788	10	US-09-758-269-3	Sequence 3, Appli
8	969	30.8	1788	10	US-09-938-842A-1444	Sequence 1444, Ap
9	938	28.8	1617	10	US-09-758-269-7	Sequence 7, Appli
10	937	28.7	1617	10	US-09-758-269-17	Sequence 17, Appli
11	522	16.6	393	10	US-09-878-574-2543	Sequence 2543, Ap
C 12	419	13.3	32679	8	US-08-976-063C-21	Sequence 1, Appli
C 13	411.5	13.1	1518	8	US-08-976-063C-21	Sequence 21, Appli
14	369	11.7	320	10	US-09-878-574-2872	Sequence 2872, Ap
C 15	346	11.0	200	9	US-09-770-696-172	Sequence 172, App
16	279	8.9	534	12	US-10-149-759-65	Sequence 65, Appli
17	260	8.3	3111	14	US-10-053-192-2	Sequence 2, Appli
18	238.5	7.6	2134	12	US-10-168-517-18	Sequence 18, Appli
19	234.5	7.4	367	9	US-09-770-791-402	Sequence 402, Appli
20	232.5	7.4	1855	12	US-10-168-517-16	Sequence 16, Appli
21	207.5	6.6	1934	12	US-10-168-517-20	Sequence 1, Appli
22	150	4.8	2037	12	US-10-168-517-1	Sequence 645, App
23	133	4.2	82993	15	US-10-080-170-645	Sequence 4052, Ap
24	127.5	4.0	13029	9	US-09-815-242-4052	Sequence 3630, Ap
25	122	3.9	5127	14	US-10-156-761-3630	Sequence 1, Appli
C 26	122	3.9	9025608	14	US-10-156-761-1	Sequence 3143, Ap
27	111.5	3.5	1875	14	US-10-156-761-3143	Sequence 4583, Ap
28	111	3.5	1956	14	US-10-156-761-4583	Sequence 1, Appli
29	110	3.5	9025608	14	US-10-156-761-1	Sequence 6784, Ap
30	108.5	3.4	4146	12	US-10-032-585-6784	Sequence 15, Appli
31	107	3.4	2825	12	US-10-289-757-15	Sequence 424, App
32	107	3.4	3075	9	US-09-833-790-424	Sequence 202, App
33	107	3.4	7515	12	US-10-101-510-202	Sequence 4, Appli
34	107	3.4	7515	12	US-10-210-120-4	Sequence 1, Appli
35	107	3.4	81940	12	US-09-759-508B-1	Sequence 724, App
36	107	3.4	81940	12	US-09-873-319-724	Sequence 1092, Ap
37	107	3.4	81940	12	US-09-960-706-1092	Sequence 7996, Ap
C 38	106.5	3.4	1083	9	US-09-815-242-7996	Sequence 82, Appli
39	106.5	3.4	10174	14	US-10-171-311-82	Sequence 76, Appli
40	106	3.4	1239	10	US-09-922-261-76	Sequence 75, Appli
41	106	3.4	1825	10	US-09-922-261-75	Sequence 2792, Ap
42	106	3.4	2514	14	US-10-156-761-2792	Sequence 1, Appli
43	106	3.4	4308	12	US-10-102-549-1	Sequence 67, Appli
44	106	3.4	4331	12	US-10-354-358-67	Sequence 3146, Ap
45	105.5	3.3	3048	14	US-10-156-761-3146	

ALIGNMENTS

RESULT 1  
US-09-758-269-5  
; Sequence 5, Application US/09758269  
; Patent No. US20020104120A1  
; GENERAL INFORMATION:  
; APPLICANT: IUCHI, SATOSHI  
; APPLICANT: KOBAYASHI, MASATOMO  
; APPLICANT: SHINOZAKI, KAZUO  
; TITLE OF INVENTION: TRANSGENIC PLANTS CARRYING NEOXANTHIN  
; TITLE OF INVENTION: CLEAVAGE ENZYME GENE  
; FILE REFERENCE: 3914-3  
; CURRENT APPLICATION NUMBER: US/09/758,269  
; CURRENT FILING DATE: 2001-01-12  
; PRIOR APPLICATION NUMBER: JP 2001-003476  
; PRIOR FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: JP 2000-010056  
; PRIOR FILING DATE: 2000-01-13  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 1800  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: CDS

i LOCATION: (1) .. (1797)

US-09-758-269-5

Alignment Scores:

Pred. No.: 0 Length: 1800  
Score: 3150.00 Matches: 599  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-09-758-269-6 (1-599) x US-09-758-269-5 (1-1800)

```
QY 1 MetAlaSerPheThrAlaThrAlaAlaValSerGlyArgTrpLeuGlyGlyAsnHisThr 20
DB 1 ATGGCTTCTTTCAACGGCAACGGCTGCGGTTTCTGGGAGATGGCTTGGTGGCAATCATACT 60

QY 21 GlnProProLeuSerSerGlnSerSerAspLeuSerTyrCysSerSerLeuProMet 40
DB 61 CAGCGCCCATATATCGTCTTCTCAAGCTCCGACTTGAGTTATTTAGTCTCTTACCTATG 120

QY 41 AlaSerArgValThrArgLysLeuAsnValSerSerAlaLeuHisThrProProAlaLeu 60
DB 121 GCCAGTGTGTACACAGTAAGTCAATGTTTCATCTGCGCTTCACACTCTCCAGCTCTT 180

QY 61 HisPheProLysGlnSerSerAsnSerProAlaIleValValLysProLysAlaLysGlu 80
DB 181 CATTTCCCTAAGCAATCATCAAACTCTCCGCCCATTTGTTAAGCCCAAGCCAAAGAA 240

QY 81 SerAsnThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaLeuAspAlaAla 100
DB 241 TCCACACATAACAGATGAATTTGTTCCAGAGCGCGCGGCGGAGCGTTTGAGCGCGCG 300

QY 101 GluGlyPheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSer 120
DB 301 GAGGTTTCTTGTTCAGCCACGAGAAGCTACACCGCTTCTCAAAACGGCTGATCTTAGT 360

QY 121 ValGlnIleAlaGlyAsnPheAlaProValAsnGluGlnProValArgArgAsnLeuPro 140
DB 361 GTTCAGATCGCGGAAATTTTGTCTCCGGTGAATGAACACCGCGCTCGCGGTAATCTTCCG 420

QY 141 ValValGlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyAlaAsnPro 160
DB 421 GTGGTCGAAAATCTCCCGATTCATCAAGAGAGTGATGTATGTCCGCAACGGAGCTTAACCCA 480

QY 161 LeuHisGluProValThrGlyHisPhePheAspGlyAspGlyMetValHisAlaVal 180
DB 481 CTTACAGAGCGGTGACAGGTCAACCATCTCTCGACGGAGAGCGGTATGGTTTACGCGCGTC 540

QY 181 LysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheVal 200
DB 541 AAATTCGAACACGGTTACGTAGCTAGCTTACGCTTTGCCGGTTTACTCAGACTAAACCGGTTTGT 600

QY 201 GlnGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGlyHis 220
DB 601 CAGGAACTCAATTTGGGTCGACCGGTTTCCCAAAGCCATCCGTTGAGCTTTCACGGCCAC 660

QY 221 ThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaAlaGlyIleValAspPro 240
DB 661 ACCGGTATTGCCGACTCATGCTATTTCTACGCCAGAGCTGCGAGCGGTATAGTCGACCCG 720

QY 241 AlaHisGlyThrGlyValAlaAsnAlaGlyLeuValTyrPheAsnGlyArgLeuLeuAla 260
DB 721 GCACACGAAACGGGTGAGTCAACCGCGTTTGGTCTATTTCAATGGCGCGGTTATGGCT 780

QY 261 MetSerGluAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysThr 280
DB 781 ATGTCGGAGGATGATTACCTTACCAAGTTTCCAGTCACTCCCAATCGAGATTTAAACACC 840

QY 281 ValGlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysVal 300
DB 841 GTTGTGCGGTTCGATTTTGTATGGACAAATTAGAAATCCCAATGATTGCCCCCAAGATC 900
```

```
QY 301 AspProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValSerLysProTyrLeu 320
DB 901 GACCCGGAATCCGGTGAACCTTTCGCTTTAAGCTACGACGTCGTTTCAAGACCTTACCTA 960

QY 321 LysTyrPheArgPheSerProAspGlyThrLysSerProAspValGluIleGlnLeuAsp 340
DB 961 AAATACCTCCGATTTCTACCGGACGGAACATAAATCACCAGGAGCTCGAGATTCAGCTTAT 1020

QY 341 GlnProThrMetMetHisAspPheAlaIleThrGluAsnPheValValValProAspGln 360
DB 1021 CAGCAACAGATGATGCACGATTTTCGATTTACAGAGAACTTCGTCGTCGTACCTGACCCAG 1080

QY 361 GlnValValPheLysLeuProGluMetIleArgGlyGlySerProValValTyrAspLys 380
DB 1081 CAAGTCGTTTTCAGCTGCGGAGATGATCCGCGTGGGTCTCCGGTGGTTTACGACAAAG 1140

QY 381 AsnLysValAlaAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSerSerAsnIleLys 400
DB 1141 AACAAAGTCGCAAGATTCGGGATTTTAGACAAATACCCGGAAGATTTCATCGAACATTAAAG 1200

QY 401 TrpIleAspAlaProAspCysPheCysPheHisLeuTrpAsnAlaTrpGluGluProGlu 420
DB 1201 TGGATTGATGCTCCAGATTTGCTTCTGCTTCCATCTCTGGAACGCTTGGGAAGACCCAGAA 1260

QY 421 ThrAspGluValValIleGlySerCysMetThrProProAspSerIlePheAsnGlu 440
DB 1261 ACAGTGAAGTCGTCGTGATAGGTCCTGTATGACTCCACGAGACTCAATTTTCAACGAG 1320

QY 441 SerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsnLeuLysThrGlyGlu 460
DB 1321 TCTGACGAGAAATCTCAAGAGTGTCTGTCTGAAATCCGCTGAATCTCAAAACCGGTGAA 1380

QY 461 SerThrArgArgProIleIleSerAsnGluAspGlnGlnValAsnLeuGluAlaGlyMet 480
DB 1381 TCAACTCGCGCTCGATCATCTCCAAACGAAAGATCAACAAGTCAACCTCGAAGCAGGGATG 1440

QY 481 ValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeuAlaLeuAlaGluPro 500
DB 1441 GTCAACAGAAACATGCTCGCGCGTAAACCAAAATTCGCTTACTTGGCTTTAGCCGAGCGG 1500

QY 501 TrpProLysValSerGlyPheAlaLysValAspLeuThrThrGlyGluValLysLysHis 520
DB 1501 TGGCTTAAGTCTCAGGATTCGCTAAAGTTGATCTCTACTCTGAGAGAGTTTAAGAACAT 1560

QY 521 LeuTyrGlyAspAsnArgTyrGlyGlyGluProLeuPheLeuProGlyGlyGlyGlu 540
DB 1561 CTTTACGCGATAACCGTTACGGAGGAGAGCCTCTGTCTTCTCCCGGAGAGGAGAGAG 1620

QY 541 GluAspGluGlyTyrIleLeuCysPheValHisAspGluLysThrTrpLysSerGluLeu 560
DB 1621 GAACAGAGAGATACATCTCTGTTTCGTTCCAGCAGAGAGACATGGAAATCCGAGTTA 1680

QY 561 GlnIleValAsnAlaValSerLeuGluValGluAlaThrValLysLeuProSerArgVal 580
DB 1681 CAGATAGTTAAACCGCTAGCTTAGAGGTTGAACGCAACGGTTAAACTTCGCTCAAGGGTT 1740

QY 581 ProTyrGlyPheHisGlyThrPheIleGlyAlaAspLeuAlaLysGlnValVal 599
DB 1741 CCGTACGATTTACCGGTACATTCATCGAGCCGATGATTTGGCGAAGACAGGTCGTG 1797
```

RESULT 2

US-09-758-269-15

; Sequence 15, Application US/09758269

; Patent No. US20020104120A1

; GENERAL INFORMATION:

; APPLICANT: IUCHI, SATOSHI

; APPLICANT: KOBAYASHI, MASATOMO

; APPLICANT: SHINOZAKI, KAZUO

; TITLE OF INVENTION: TRANSGENIC PLANTS CARRYING NEOXANTHIN

; FILE REFERENCE: 3914-3

; CURRENT APPLICATION NUMBER: US/09/758,269

; CURRENT FILING DATE: 2001-01-12







Db 1546 TGCCCCAAAGTCGCGGCTTTGGAAAGTTGATTTGCTGAGTGGGAAGTGAAGAAGTAC 1605  
Qy  
Db 521 LeuTyrGlyAspAsnArgTyrGlyGluProLeuPheLeuProGlyGluGlyGlyGlu 540  
Qy  
Db 1606 ATGTATGGAGAGAGAAAGTTCGGTGGGAGCCTCTGTTCTTCCC---AACGCCCAAAA 1662  
Qy  
Db 541 GluAspGluGlyTyrIleLeuCysPheValHisAspGluLysThrTrpLysSerGluLeu 560  
Qy  
Db 1663 GAAGACGATGGGTATATTCTGGCATTCGTGCACGACGAGAAAGAAATGGAATCCGAGCTG 1722  
Qy  
Db 561 GlnIleValAsnAlaValSerLeuGluValGluAlaThrValLysLeuProSerArgVal 580  
Qy  
Db 1723 CAGATTGGAATGCCCAAAATTTAAAGCTCGAAGCTTCATCAAACTCCCTCTCGTGT 1782  
Qy  
Db 581 ProTyrGlyPheHisGlyThrPheIleGlyAlaAspAspLeuAlaLysGln 597  
Qy  
Db 1783 CCTACGGTTTTCATCGAACTTTCATTCATCCAGGATTTGAGGAACAA 1833

## RESULT 4

US-09-758-269-1  
; Sequence 1, Application US/09758269  
; Patent No. US20020104120A1  
; GENERAL INFORMATION:  
; APPLICANT: IUCHI, SATOSHI  
; APPLICANT: KOBAYASHI, MASATOMO  
; APPLICANT: SHINOZAKI, KAZUO  
; TITLE OF INVENTION: TRANSGENIC PLANTS CARRYING NEOXANTHIN  
; TITLE OF INVENTION: CLEAVAGE ENZYME GENE  
; FILE REFERENCE: 3914-3  
; CURRENT APPLICATION NUMBER: US/09/758,269  
; PRIOR FILING DATE: 2001-01-12  
; PRIOR APPLICATION NUMBER: JP 2001-003476  
; PRIOR FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: JP 2000-010056  
; PRIOR FILING DATE: 2000-01-13  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 1752  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1749)  
US-09-758-269-1

## Alignment Scores:

Pred. No.: 9.18e-219 Length: 1752  
Score: 1991.00 Matches: 380  
Percent Similarity: 78.44% Conservative: 82  
Best Local Similarity: 64.52% Mismatches: 105  
Query Match: 63.21% Indels: 22  
DB: 10 Gaps: 6

US-09-758-269-6 (1-599) x US-09-758-269-1 (1-1752)

Qy 16 GlyGlyAsnHisThrGlnProLeuSerSerGlnSerSerAspLeuSerTyr--- 34  
Db 31 GGTGGTATTAAACATGGCCT-----CAGGCCAAATTTGATTTGGTATTAGG 78  
Qy 35 ---CysSerSerLeuProMetAlaSerArgValThrArgLysLeuAsnValSerSerAla 53  
Db 79 CCATTAAAGAACCAACGAAGGTATTAAATGCACGGTGCAGATCGACGTAACGGAA--- 135  
Qy 54 LeuHisThrProProAlaLeuHisPheProLysGlnSerSerAsnSerProAlaIleVal 73  
Db 136 TTAACCAAAAAACGCGCAATTTATTACACCCAGAACCCGCTACTCCGCGC----- 186  
Qy 74 ValLysProLysAlaLysGluSerAsnThrLysGlnMetAsnPheGlnArgAlaAla 93  
Db 187 -----CAGCATATCTCTCCGCGTAAACATCTTCAGAAAGCGCG 228

Qy 94 AlaAlaLeuLeuAspAlaAlaGluGlyPheLeuValSerHisGluLysLeuHisProLeu 113  
Db 229 GCGATTGGATCGACGCGCTGAGCGTGCATTAATCTCACAGCAAGATTTCTCCACTT 288  
Qy 114 ProLysThrAlaAspProSerValGlnIleAlaGlyAsnPheAlaProValAsnGluGln 133  
Db 289 CCAAAACCGCTGATCCAGCTGTTCCAGATTGCGCGGAATTATTCGCCCGTACCGGAATCT 348  
Qy 134 ProValArgArgAsnLeuProValValGlyLysLeuProAspSerIleLysGlyValTyr 153  
Db 349 TCCGTCGCGGGAACCTCACCGTCGAAGGAACAATCCCTGACTGCTGATTCACCGGTGTAT 408  
Qy 154 ValArgAsnGlyAlaAsnProLeuHisGluProValThrGlyHisPhePheAspGly 173  
Db 409 ATCCGTAACGCGCGCAATCCGATGTTGAGCCACACAGCTGGGCACCATTTATTCACGGA 468  
Qy 174 AspGlyMetValHisAlaValLysPheGluHisGlySerAlaSerTyrAlaCysArgPhe 193  
Db 469 GACGGAATGGTTTCACGCAAGTTAAACACCGGTTGAGTGTAGCTAGCTACGCGGT 528  
Qy 194 ThrGlnThrAsnArgPheValGlnGluArgGlnLeuGlyArgProValPheProLysAla 213  
Db 529 ACAAAACCGAGAGATTGGTTCAGAAACCAATGGTGGTGCACCACTTTCCCGAAGCA 588  
Qy 214 IleGlyGluLeuHisGlyHisThrGlyIleAlaArgLeuMetLeuPheTyrAlaArgAla 233  
Db 589 ATCGCGGAGCTTCACGGTCACTCGGGAATCGCACGCTTGTGCTGTTTACGCACGTGG 648  
Qy 234 AlaAlaGlyIleValAspProAlaHisGlyThrGlyValAlaAsnAlaGlyLeuValTyr 253  
Db 649 CTTTGTGTCTGATCAACCAACCAACCGCGCTCGGAGTAGCAACCGCGTTTGGTTTAC 708  
Qy 254 PheAsnGlyArgLeuLeuAlaMetSerGluAspAspLeuProTyrGlnValGlnIleThr 273  
Db 709 TTTATACCGCGCTTTTAGCTATGTACAGACGATTTACCGTACCAATTAATAATTACT 768  
Qy 274 ProAsnGlyAspLeuLysThrValGlyArgPheAspPheAspGlyGlnLeuGluSerThr 293  
Db 769 CAACCGCGGATCTCAACACCGTTGGACGTTACGATTTTCGACGGTCAGTTAAATCCGCA 828  
Qy 294 MetIleAlaHisProLysValAspProGluSerGlyGluLeuPheAlaLeuSerTyrAsp 313  
Db 829 ATGATAGCTACCCCGAAACTGGACCGGTTACGAAGAGAGCTTCACGCGTTAAGTACGAC 888  
Qy 314 ValValSerLysProTyrLeuLysTyrPheArgPheSerProAspGlyThrLysSerPro 333  
Db 889 GTCGTTAAGAAACCTTACTGAATATCTCAGATTCTCCAGACGCGGTAAATCCGCG 948  
Qy 334 AspValGluIleGlnLeuAspGlnProThrMetHisAspPheAlaIleThrGluAsn 353  
Db 949 GAATTGGAGATCCCGCTCGAAACTCCGACGATGATTACGATTTTCGCTATTAACGGAGAAT 1008  
Qy 354 PheValValProAspGlnGlnValValPheLysLeuProGluMetIleArgGlyGly 373  
Db 1009 TTTGTGGTGAATTCCTGTGATCAACAACTCGTGTTCAGCTCGCGAGATGATTTCCGCTAAA 1068  
Qy 374 SerProValValTyrAspLysAsnLysValAlaArgPheGlyIleLeuAspLysTyrAla 393  
Db 1069 TCTCCGGTGTGTTTCGACGGAGAAAGGTTTCCCGATTTGGGATTAATCCCAAGGACGCG 1128  
Qy 394 GluAspSerSerAsnIleLysTrpIleAspAlaProAspCysPheCysPheHisLeuTyr 413  
Db 1129 ACAGAAGCTTCTCAGATAATCTGGGTGAATCTCCGGAGACGCTTCTGTTTTCATCTCTGG 1188  
Qy 414 AsnAlaTrpGluProGluThrAspGluValValIleGlySerCysMetThrPro 433  
Db 1189 AATGATCGGAATCCCGGACGAGAGAGATTTGTGGTATCGGATCGTGTATGTGCGCG 1248  
Qy 434 ProAspSerIlePheAsnGluSerAspGluAsnLeuLysSerValLeuSerGluIleArg 453  
Db 1249 GCGGATTCATCTTCAACGAGAGACGAGAGCTTTGAGAAGCGTTTTCGAGATCAGG 1308  
Qy 454 LeuAsnLeuLysThrGlyGluSerThrArgArgProIleIleSerAsnGluAspGlnGln 473

```
Db 1309 ATAAACCTCAGACACGTAACACACCGCTCGTTGGTTGGTTAAACGAGAT----- 1362
Qy 474 ValAsnLeuGluAlaGlyMetValAsnArgAsnMetLeuGlyArgLysThrLysPheAla 493
Db 1363 GTAAATTTAGAGATTGGTATGTTAAACCGGAACCGGTTAGGAAGAAAAACCGGTTCCGC 1422
Qy 494 TyrLeuAlaLeuAlaGluProTrpProLysValSerGlyPheAlaLysValAspLeuThr 513
Db 1423 TTTTGGCTATGCTTATCTTGGCCAAAGTTTCCCGGTTTCGGCTTAAGGTGATCTTTGC 1482
Qy 514 ThrGlyGluValLysLysHisLeuTyrGlyAspAsnArgTyrGlyGlyGluProLeuPhe 533
Db 1483 ACCGGTAGATGAAAAAATATATTACCGCGTGAGAAATATGGCGGCAACCGGTTTTC 1542
Qy 534 LeuProGlyGluGlyGly-----GluGluAspGluGlyTyrIleLeuCysPhe 549
Db 1543 TTGCCCGCAACTCCGGTAACGGCGGAAGAAATGAAGATGACGGTTATATATTGTCAC 1602
Qy 550 ValHisAspGluLysThrTrpLysSerGluLeuGlnIleValAsnAlaValSerLeuGlu 569
Db 1603 GTTCATGACGAAGAAACAAACACATCAGAGCTTCAGATTATTAACGCTGTTAAATTTAAG 1662
Qy 570 ValGluAlaThrValLysLeuProSerArgValProTyrGlyPheHisGlyThrPheIle 589
Db 1663 CTTGAAGCTACGATTAAACTACCGCTAGAGTACCGTATGCGTTTCATGGCACATTGTCG 1722
Qy 590 GlyAlaAspLeuAlaLysGlnVal 598
Db 1723 GATTGGAATGAACCTCGTTGATCAATTA 1749
```

## RESULT 5

```
US-09-758-269-13
; Sequence 13, Application US/09758269
; Patent No. US20020104120A1
; GENERAL INFORMATION:
; APPLICANT: IUCHI, SATOSHI
; APPLICANT: KOBAYASHI, MASATOMO
; APPLICANT: SHINOZAKI, KAZUO
; TITLE OF INVENTION: TRANSGENIC PLANTS CARRYING NEOXANTHIN
; FILE REFERENCE: 3914-3
; CURRENT APPLICATION NUMBER: US/09/758,269
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: JP 2001-003476
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: JP 2000-010056
; PRIOR FILING DATE: 2000-01-13
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 1815
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1812)
US-09-758-269-13
```

## Alignment Scores:

Pred. No.:	1,03e-211	Length:	1815
Score:	1930.00	Matches:	375
Percent Similarity:	75.68%	Conservative:	73
Best Local Similarity:	63.34%	Mismatches:	122
Query Match:	61.27%	Indels:	22
DB:	10	Gaps:	8

US-09-758-269-6 (1-599) x US-09-758-269-13 (1-1815)

```
Qy 22 ProProLeuSerSerSerGlnSerSerAspLeuSerTyrCysSerSerLeuProMetAla 41
Db 49 CCGGCCCGGTCAGGCGCGGCGCTCC-----AATTCGTCAGGTTCTCG 93
```

```
Qy 42 SerArgValThrArgLysLeuAsnValSerSerAlaLeuHisThrPro-----Pro 58
Db 94 CCGCGCCCGTCAGCTCGGTGCGCGCCGCGAGTGCCTCCAGGCGCGCTTCCACAAGCCC 153
Qy 59 AlaLeuHisPheProLysGlnSerSerAsnSerProAlaIleValVal-----74
Db 154 GTCCGCCACCTGCTCGCGCTCCAGGAAGCCCGCCCATTTGCGTCCAGGCGACGCC 213
Qy 75 ---LysProLysAlaLysGluSerAsnThrLysGlnMetAsnLeuPheGlnArgAlaAla 93
Db 214 CCGCGCGCGAAGGAGCGGCGGCGCAAGAAGCAGCTCAACTTGTTCAGCGCGCGCG 273
Qy 94 AlaAlaAlaLeuAspAla---AlaGluGlyPheLeuValSer---HisGluLysLeuHis 111
Db 274 CCGCGCCCGCTCGACGCGTTTCGAGGAAGGGTTTCGTGCCCAACGTCCTCGAGCGGCCAC 333
Qy 112 ProLeuProLysThrAlaAspProSerValGlnIleAlaGlyAsnPheAlaProValAsn 131
Db 334 GGGCTGCCAGCAGCGCGCGCGTGCAGATCGCGCAACTTCGCGCCCGTCGGG 393
Qy 132 GluGlnProValArgAsnLeuProValValGlyLysLeuProAspSerIleLysGly 151
Db 394 GAGAGCGCGCGCTGTCAGAGCTCCCGCTCTCCGCGCGCATCCCGCCCTTCATCGAGCGG 453
Qy 152 ValTyrValArgAsnGlyAlaAsnProLeuHisGluProValThrGlyHisHisPhePhe 171
Db 454 GTCTACCGCGCAACGGCGCAACCCCTCTGACCCCGCTCGCGGGGCGCACCTCTTC 513
Qy 172 AspGlyAspGlyMetValHisAlaValLysPheGluHisGlySerAla---SerTyrAla 190
Db 514 GAGCGCAGCGCATGGTGTGACGCGCTCGGATACGCAACGCGCGCGCGAGTCTCTACGCC 573
Qy 191 CysArgPheThrGlnThrAsnArgPheValGlnGluArgGlnLeuGlyArgProValPhe 210
Db 574 TGCCTCTTCAGGAGACGCGCGCTCGCGCAGAGCGCGCATCGCGCGCGCTCTTC 633
Qy 211 ProLysAlaIleGlyGluLeuHisGlyHisThrGlyIleAlaArgLeuMetLeuPheTyr 230
Db 634 CCCAAGGCCATTGGCGAGCTGCACGGGACCTCCCGGATCGCGCGCTTGCCTCTTCTAC 693
Qy 231 AlaArgAlaAlaGlyIleValAspProAlaHisGlyThrGlyValAlaAlaAsnAlaGly 250
Db 694 GCGCGCGCGCTCGCGCTCGTGACCCCTCGCGCGCGCACCGCGCGTGGCCACCGCGGC 753
Qy 251 LeuValTyrPheAsnGlyArgLeuAlaMetSerGluAspAspLeuProTyrGlnVal 270
Db 754 CTCGTCTACTTCAACGGCGCGCTGCTCGCCATGTTCGAGGACGACCTCCCTACACCGTC 813
Qy 271 GlnIleThrProAsnGlyAspLeuLysThrValGlyArgPheAspPheAspGlyGlnLeu 290
Db 814 CCGTGGGAGACGCGCGGACCTCGAGACCGTCGCGCGCTACGACTTCGACGGGCGAGCTC 873
Qy 291 GluSerThrMetIleAlaHisProLysValAspProGluSerGlyGluLeuPheAlaLeu 310
Db 874 GGCTGCGCCATGATCGCGCACCCCAAGCTGAGCCCGCGCACCGGGAGCTCCACGCGCTC 933
Qy 311 SerTyrAspValValSerLysProTyrLeuLysTyrPheArgPheSerProAspGlyThr 330
Db 934 AGCTACGACGTCATCAAGAGCGCGTACTCTCAAGTACTTCTACTTCAGCGCGCGCGCAC 993
Qy 331 LysSerProAspValGluIleGlnLeuAspGlnProThrMetHisAspPheAlaIle 350
Db 994 AAGTCCGACGAGTGGAGATCCCGCTGGAGCAGCCCGAGATGATCCAGCATTCGCCATC 1053
Qy 351 ThrGluAsnPheValValProAspGlnGlnValValPheLysLeuProGluMetIle 370
Db 1054 ACCGAGAACTTCGTGGTTGTGCCCGCACCCAGGTGGTGTTCAGCTCCAGGAGATGCTG 1113
Qy 371 ArgGlyLysSerProValValTyrAspLysAsnLysValAlaArgPheGlyIleLeuAsp 390
Db 1114 CCGCGCGCGTCCCGCGGTGCTCGAAGAGAGACGTCGCGGTTTCGGGCTGCTCCCC 1173
Qy 391 LysTyrAlaGluAspSerSerAsnIleLysTrpIleAspAlaProAspCysPheCysPhe 410
```

```

1174 AACGACCGCGGACGCGTGGAGATGGCGTGGAGCTGCGGACTGCTCTCTTC 1233
411 HisLeuTrpAsnAlaTrpGluGluProGluThrAspGluValValValIleGlySerCys 430
1234 CACCTGTGGAACGCGTGGAGGACGAGCGGCGGAGGTGGTGTGTGCTCTCTGC 1293
431 MetThrProProAspSerIlePheAsnGluSerAspGluAsnLeuLysSerValLeuSer 450
1294 ATGACCGCGCGGACCTCATCTTCAACGAGTCCGAGCGCGCTGGAGAGCGTGTGACC 1353
451 GluIleArgLeuAsnLeuLysThrGlyGluSerThrArgArgProIleIleSerAsnGlu 470
1354 GAGATCCGCTGGACCGCGGACGCGGCGGCTCCACGCGCGCGCTCTCTGCGG---CCG 1410
471 AspGlnGlnValAsnLeuGluAlaGlyMetValAsnArgAsnMetLeuGlyArgLysThr 490
1411 TCGCAGCAGGAGAACCTGGAGGTGGCATGGTGAACCGCAACTGCTGGCGCGGAGACC 1470
491 LysPheAlaTyrlleuAlaLeuAlaGluProTrpProLysValSerGlyPheAlaLysVal 510
1471 CGGTACGCGTACCTCGCGGTGGCGGAGCGGTGGCCCAAGGAGTTCGGGCTTCGCAAGGAG 1530
511 AspLeuThrThrGlyGluValLysLysHisLeuTyrlGlyAspAsnArgTyrlGlyGlu 530
1531 GACCTGTCCAGCGCGGAGCTCACCAGTTCGAGTACGGCGGCGGCGGTTCGGCGCGGAG 1590
531 ProLeuPheLeuProGlyGluGlyGlyGlu-----GluAspGluGlyTyrl 545
1591 CCCTGCTGTTCCCATGACCGCGCGGCGGCGGCGGCGGCGGCGGCGGAGACGCGGTAC 1650
546 IleLeuCysPheValHisAspGluLysThrTrpLysSerGluLeuGlnIleValAsnAla 565
1651 GTGCTCACCTTCGTCACGACGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1710
566 ValSerLeuGluValGluAlaThrValLysLeuProSerArgValProTyrlGlyPheHis 585
1711 GCGGACATCCGCTGGAGGCCACGGTTCAGTCCGCGTCCGCGTCCGCGTCCGCTTCGCCAC 1770
586 GlyThrPheIleGlyAlaAspLeuAlaLysGln 597
1771 GGCACCTTCATCAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1806

```

## RESULT 6

```

US-09-758-269-9
; Sequence 9, Application US/09/58269
; Patent No. US20020104120A1
; GENERAL INFORMATION:
; APPLICANT: IUCHI, SATOSHI
; APPLICANT: KOBAYASHI, MASATOMO
; APPLICANT: SHINOZAKI, KAZUO
; TITLE OF INVENTION: TRANSGENIC PLANTS CARRYING NEOXANTHIN
; TITLE OF INVENTION: CLEAVAGE ENZYME GENE
; FILE REFERENCE: 3914-3
; CURRENT APPLICATION NUMBER: US/09/758,269
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: JP 2001-003476
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: JP 2000-010056
; PRIOR FILING DATE: 2000-01-13
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1731)
US-09-758-269-9

```

Alignment Scores:

Pred. No.: 4.83e-181 Length: 1734

```

Score: 1663.50 Matches: 323
Percent Similarity: 72.56% Conservative: 100
Best Local Similarity: 55.40% Mismatches: 137
Query Match: 52.81% Indels: 23
DB: 10 Gaps: 9

```

US-09-758-269-6 (1-599) x US-09-758-269-9 (1-1734)

```

QY 24 LeuSerSerGlnSerSerAspLeuSerTyrlCysSerSerLeuProMetAlaSerArg 43
DB 25 CTTCTCCGACGACGACITCTCTCTCGTCTCATTTTACTTCCACACCCCAAAATGCAAT 84
QY 44 ValThrArgLysLeuAsnValSerSer---AlaLeuHisThrProProAlaLeuHisPhe 62
DB 85 ATTCTCGACGAATTTCTCATTAACCTTTCAAGATACCGACACTTCTCTGATCTCACTTCT 144
QY 63 ProLysGlnSerSerAsnSerProAlaIleValValLysProLysAlaLysGluSerAsn 82
DB 145 CCGGTTCGCTCACCG-----GTTAAGCTCAAACCAACGATATCCAAAC 186
QY 83 ThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaLeuAspAlaAlaGlu--- 101
DB 187 -----TTAAACCTTCTTCAGAGCTAGCGGTACGATCTCGACCAAGATTGAGTCC 237
QY 102 GlyPheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSerVal 121
DB 238 TCTATCGTTATTCCTATGAGCAGAATCGCGCTTCTCTAAACCGACCGACCGCGGTT 297
QY 122 GlnIleAlaGlyAsnPheAlaProValAsnGluGlnProValArgAsnLeuProVal 141
DB 298 CAATTATCAGGTAACTTCGCTCCGTTAATGAATGTCGGTTTCAGAACGGTTTAGAAGTG 357
QY 142 ValGlyLysLeuProAspSerIleLysGlyValTyrlValArgAsnGlyValAlaAsnProLeu 161
DB 358 GTTGTGTCAGATTCCTTCTGTCTAAAGAGATTACATCCGTAAACCGTCAACACCTATG 417
QY 162 HisGluProValThrGlyHisPhePheAspGlyAspGlyMetValHisAlaValLys 181
DB 418 TTTCGCGGTTAGCGGACATCATTTATTTACGCGTGACGGAATGATTACACCGCTTACT 477
QY 182 -----PheGluHisGlySerAlaSerTyrlAlaCysArgPheThrGlnThrAsnArgPhe 199
DB 478 ATCGGTTTGTATAAC---CAGGTTAGTTACAGCTCCGCTACACTTAAACCAACACCGGTT 534
QY 200 ValGlnGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGly 219
DB 535 GTTCAAGAAACCGCGCTTGGACGATCGGTTTCCCTAAACCAATCCGCGAGGTTCCAGGC 594
QY 220 HisThrGlyIleAlaArgLeuMetLeuPheTyrlAlaArgAlaAlaAlaGlyIleValAsp 239
DB 595 CATTCCGGTCTAGCTCGACTCGCTCTCTTACGCGCTCGAGCTGGGATCGGTCTAGTGGAC 654
QY 240 ProAlaHisGlyThrGlyValAlaAsnAlaGlyLeuValTyrlPheAsnGlyArgLeuLeu 259
DB 655 GGGACACGTCGTCGGCGTAGTAACCGCGGTGTGGTTTCTTTAAACCGGCAAGTTATTA 714
QY 260 AlaMetSerGluAspAspLeuProTyrlGlnValGlnIleThrProAsnGlyAspLeuLys 279
DB 715 GCCATGTCGAAGATGATCTTCTTACCAGTAGTGAAGATCGAGCTGCAAGGAGATCTTGAG 774
QY 280 ThrValGlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLys 299
DB 775 ACGATCGGACGCTCGGATTCGATCCAGCAGATTTGACTCTTCAGTATAGCGCATCTTAAG 834
QY 300 ValAspProGluSerGlyGluLeuPheAlaLeuSerTyrlAspValValSerLysProTyrl 319
DB 835 GTGGACGCGCACGACGAGGATCTCCATACACTAGCTACAACGTTTGAAGAAACCTCAT 894
QY 320 LeuLysTyrlPheArgPheSerProAspGlyThrLysSerProAspValGluIleGlnLeu 339
DB 895 CTCAGGTATCTTAATTCACACGTCGCGGAAAAGACACGTCGCTGGAGATCACGCTC 954
QY 340 AspGlnProThrMetMetHisAspPheAlaIleThrGluAsnPheValValProAsp 359

```

```
Db 955 CCTGAACACGATGATTCATGATTTCCGATACCGAGAAATTTGTGTTATACCGGAT 1014
Qy 360 GlnGlnValValPheLysLeuProGluMetIleArgGlySerProValValTyrAsp 379
Db 1015 CAGCAATATCGTATTCAAAATATCCGAATGATTCGGGCGGGTCAACCGTTTACGTT 1074
Qy 380 LysAsnLysValAlaAtrPheGlyIleLeuAspLysTyrAlaGluAAspSerSerAsnIle 399
Db 1075 AAAGAAAATATGCGAGATTTGGAGTTTGTCAAGCAGGATCTGACCGGTCCGATATA 1134
Qy 400 LysTrpIleAspAlaProAspCysPheCysPheHisLeuTrpAsnAlaTrpGluGluPro 419
Db 1135 AATTGGGTGATGATACCGGATGTTCTGTTCCATCTATGGAATGCGTGGGAAGAG-- 1191
Qy 420 GluThrAspGlu-----ValValIleGlySerCysMetThrProProAsp 435
Db 1192 AGAACCGAAGAGGAGACCCAGTTATCTGCTGAATCGGTTCATGATGAGCCACCCGAC 1251
Qy 436 SerIlePheAsnGluSerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsn 455
Db 1252 ACGATCTTTAGTGAATCAGGAACCAACCCGGGTGGAATTAAGTGAGATCCGGTTAAAC 1311
Qy 456 LeuLysThrGlyGluSerThrArgArgProIleIleSerAsnGluAAspGlnGlnValAsn 475
Db 1312 ATGGGTACAAAGAAATCGAACGTAAGTTATCGTAACCGGA-----GTGNAAT 1359
Qy 476 LeuGluAlaGlyMetValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeu 495
Db 1360 TTAGAAGCGGGTCACATAAACCGTAGTTACGTGGGCGGAAAGCCAGTTCGTTTACATA 1419
Qy 496 AlaLeuAlaGluProTrpProLysValSerGlyPheAlaLysValAspLeuThrThrGly 515
Db 1420 GCAATAGCCGATCTCTGGCCCAATATGCGAGTGGCATTCGGAAGGTAGATATACAAACGGC 1479
Qy 516 GluValLysLysHisLeuTyrGlyAspAsnArgTyrGlyGlyGluProLeuPheLeuPro 535
Db 1480 ACCGTTTCAGAGTTTAATACGGACCGAGCGGTTCGGTGGGAACCGTCTTGTACCG 1539
Qy 536 GlyGluGlyGlyGluGluAspGluGlyTyrIleLeuCysPheValHisAspGluLysThr 555
Db 1540 GAGGAGAAGGAGAAGAAAGAGGTTATGTAATGGGGTTTGTGAGAGACCAAGAGAAA 1599
Qy 556 TrpLysSerGluLeuGlnIleValAsnAlaValSerLeuGluValGluAlaThrValLys 575
Db 1600 GACGAGTCGGAGTTTGTGGTGTGACGCGACGGATATGAAGCAAGTCGCGCGGTGCGC 1659
Qy 576 LeuProSerArgValProTyrGlyPheHisGlyThrPheIleGlyAlaAspLeuAla 595
Db 1660 TTGCGGAGAGGGTACCTTATGTTTCCATGGAACGTTTCGTGAGCGAGAAATCAGTTGAAG 1719
Qy 596 LysGlnVal 598
Db 1720 GAACAAGTT 1728
```

## RESULT 7

```
US-09-758-269-3
; Sequence 3, Application US/09758269
; Patent No. US20020104120A1
; GENERAL INFORMATION:
; APPLICANT: TUCHI, SATOSHI
; APPLICANT: KOBAYASHI, MASATOMO
; APPLICANT: SHINOZAKI, KAZUO
; TITLE OF INVENTION: TRANSGENIC PLANTS CARRYING NEOXANTHIN
; TITLE OF INVENTION: CLEAVAGE ENZYME GENE
; FILE REFERENCE: 3914-3
; CURRENT FILING DATE: 2001-01-12
; PRIOR FILING DATE: 2001-01-12
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: JP 2000-003476
; PRIOR APPLICATION NUMBER: JP 2000-010056
; PRIOR FILING DATE: 2000-01-13
; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1788
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1785)
US-09-758-269-3
Alignment Scores:
Pred. No.: 5,15e-101 Length: 1788
Score: 969.00 Matches: 214
Percent Similarity: 55.31% Conservative: 114
Best Local Similarity: 36.09% Mismatches: 221
Query Match: 30.76% Indels: 44
DB: 10 Gaps: 13
US-09-758-269-6 (1-599) x US-09-758-269-3 (1-1788)
Qy 32 LeuSerTyrCysSerSerLeuProMetAlaSerArgValThrArgLysLeuAsnValSer 51
Db 56 CTCTTCTTCGCGCGGATCTTCTCTC-----CTACTC 88
Qy 52 SerAlaLeuHisThrProAlaLeuHisPheProLysGlnSer-----SerAsn 68
Db 89 TCCTCCGTATCAACTCCGCGTCTCGAAGACGTTCTCCATCACAACCCAGCGACA 148
Qy 69 SerProAlaIleValLysProLys-AlaLysGluSerAsnThrLysGlnMetAsnLe 88
Db 149 ACAATGATCGTCGTCAAAACCCCAACACCTCCACAAACCAACCAATCACAACCTTAGTCT 208
Qy 88 uPheGlnArgAlaAlaAlaAlaLeuAspAla-AlaGluGlyPheLeuValSerHisG 108
Db 209 CATCACCCAGAACTCCGACCAAGAAATGACTCTCGCAACAGCTCTTTCACCCCGTCG 268
Qy 108 LuLysLeu-----HisProLeuProLysThrAlaAspProSerValG 122
Db 269 AAGATGTAATCAACACGTTTCATCGATCCACTTCACGTCCTTCCTGTTGATCCAAAACATG 328
Qy 122 InIleAlaGlyAsnPheAlaProValAsnGlnProValArgArgAsnLeuProValV 142
Db 329 TCCTCTCTGATAACTTCGCTCTCTCGACGAGTTCCTCCAAACAGACTGTGAAATCA 388
Qy 142 al---GlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyAlaAsnProL 161
Db 389 TCCACGCGACTCTTCCACTGTAAACGCGCTTATCATCCGTAACGGTCCAAATCCAC 448
Qy 161 euHisGluProValThrGlyHisPhePheAspGlyAspGlyValHisAlaValL 181
Db 449 AGTTTCTCCCTCGTGGTCTTACCATCTCTTCGACGGGCGAGGTATGCTTCACGCCATAA 508
Qy 181 ysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheValG 201
Db 509 AAATCCCAACCGGTAAAGCCACTCTCTGTAGCAGATACGTCAAGACTTATAATACACAG 568
Qy 201 InGluArgGlnLeuArgProValPheProLysAlaIleGlyGluLeuHisGlyHisT 221
Db 569 TCGAAGAAACAAACCGGAGCTCCGTTATGCTTAACGTTTCCCGGATTCACCGGTGTA 628
Qy 221 hr---GlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaAlaGlyIleValAspP 240
Db 629 CGGCGTCAGTAGCTCGTGGAGCTTTAAGCGGAGCTAGGTTTAAACCGGACAGTATAATC 688
Qy 240 roAlaHisGlyThrGlyValAlaAlaAsnAlaGlyLeuValTyrPheAsnGlyArgLeuLeuA 260
Db 689 CGGTTAACGGCATTTGGTTTAGCTAATACAAGCTTAGCTTTCTCAGTAACCGTCTCTTTG 748
Qy 260 laMetSerGluAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysT 280
Db 749 CTTTAGGTGATCTGATTTACCTACCGCGTCCGCTTAACCGAATCAGGAGATATGAAA 808
Qy 280 hrValGlyArgPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysV 300
```

```
Db 809 CGATCGACGGTACGATTTTCGACGGAAATTCGATGAGTATGACAGCTCATCCCTAAAA 868
Qy 300 aAspProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValSerLysProTyrL 320
Db 869 CCGATCCAAATACCGGAGAACTTCGCTTCGGGTACGGTCCGGTT---CCACCGTTT 925
Qy 320 euLysTyrPheArgPheSerProAspGlyThrLysSerProAspValGluLeu---GlnL 339
Db 926 TACATATTTTCGGTTTGAATCCCGCGGAAACAAAGAGACGCTTCGATATTCCTCGA 985
Qy 339 euAspGlnProThrMetHisAspPheAlaLeuThrGluAsnPheValValProA 359
Db 986 TGACGCTCGCTGCTTCCTCATGACTTCGGATCACGAAACGCTACGGGATTTTCGCGAG 1045
Qy 359 spGlnGlnValValPheLys-----LeuProGluMetIleArgGlyGlySerProV 376
Db 1046 AGATTCCAGCTGTCGATGAGGATGACATGTTGGATTTGTTCTCGAAGGTGGTTCTCCG 1105
Qy 376 alValTyrAspLysAsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAsp 396
Db 1106 TTGGTACTGATACCGGAAACCTCCAGGCTTCGAGTGATTCCTAAGTACGCGCGGAGATG 1165
Qy 396 erSerAsnIleLysTrpIleAspAlaProAspCysPheCysPheHisLeuTrpAsnAla 416
Db 1166 AGTCGAGATGAAATGGTTCGAAGTTCCTGGATTCATATCATTCACGCTATTAATGCTT 1225
Qy 416 rpGluGluProGluThrAspGluValValIleGlySerCysMetThrProProAsp 436
Db 1226 GGGATCAAGATGATGAAACAGCGCTGTTTGGATTCACCGCAATATATGTCGATGAA 1285
Qy 436 erIlePheAsnGluSerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsnL 456
Db 1286 ATACTTTAGAGAGATGAT---CTGGTTCATGCTTTGGTGGAGAGGTGAAGATCGATC 1342
Qy 456 euLysThrGlyGluSerThrArgArgProIleIleSerAsnGluAspGlnGlnValAsnL 476
Db 1343 TCGTACCGGGATTGTGAGACGTCATCCGATCTCAGCGAGG-----AATC 1387
Qy 476 euGluAlaGlyMetValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeuA 496
Db 1388 TCGATTTCCGCTGTGATTATCCGCGCTTCTCGGAGATGTAGCAGGTACGTTTACGCG 1447
Qy 496 laLeuAlaGluProTrpProLysValSerGlyPheAlaLysValAspLeuThrThrGly 516
Db 1448 CGATTGGAGATCCGATGCGCAAGATCTCCGGTGTGGTGAAGCTGTATGTCTAAAGGAG 1507
Qy 516 lu-----ValLysLysHisLeuTyrGlyAspAsnArgTyrGlyGluP 531
Db 1508 ATCGGATGATTGTACGGTGGCCCGGAGAAATGTACGGTTTCAGGTTGTTACGCGCGGAG 1567
Qy 531 roLeuPheLeuProGlyGlyGlyGly-----GluGluAspGluGlyTyrIleL 547
Db 1568 CGTTTTCGTAGTAGGATCTGGTAATCCGGAGCGCGAGAGAGATGATGTTATGTGG 1627
Qy 547 euCysPheValHisAspGluLysThrTrpLysSerGluLeuGlnIleValAsnAlaVal 567
Db 1628 TGACGTATGTTCCAGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1687
Qy 567 er-----LeuGluValGluAlaThrValLysLeuProSerArgValProTyrGlyPheH 585
Db 1688 CCGCGGAGCTTGAATTCGCGCGCGCGGAGGTGCGCGAAGGGTTCGCTACGATTC 1747
Qy 585 isGlyThrPheIleGlyAlaAspAspLeuAlaLys 596
Db 1748 ATGGGTATTGTCAAGGAAAGTACCTTAATAAG 1782
```

## RESULT 8

```
US-09-938-842A-1444
; Sequence 1444, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
```

```
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1444
; LENGTH: 1788
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-09-938-842A-1444

Alignment Scores:
Pred. No.: 5,15e-101 Length: 1788
Score: 969.00 Matches: 214
Percent Similarity: 55.31% Conservative: 114
Best Local Similarity: 36.09% Mismatches: 221
Query Match: 30.76% Indels: 44
DB: 10 Gaps: 13

US-09-758-269-6 (1-599) x US-09-938-842A-1444 (1-1788)
```

```
Qy 32 LeuSerTyrCysSerSerLeuProMetAlaSerArgValThrArgLysLeuAsnValSer 51
Db 56 CTCTTCTTCGCGCGGATCTCTCTCTC-----CTACTC 88
Qy 52 SerAlaLeuHisThrProProAlaLeuHisPheProLysGlnSer-----SerAsn 68
Db 89 TCTCTCGTATCAACTCCGCGCTCGAAGAACGTTCTTCCAATCAAAACCAACGAGCA 148
Qy 69 SerProAlaIleValValLysProLysAlaLysGluSerAsnThrLysGlnMetAsnLe 88
Db 149 ACAATGATCGTGTAAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 208
Qy 88 uPheGlnArgAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 108
Db 209 CATCACCCGAAATCTCCGACGAGAAATGACTCTCGCAACAGCTCTCTTCCACCGTCG 268
Qy 108 luLysLeu-----HisProLeuProLysThrAlaAspProSerValG 122
Db 269 AAGATGTAATCAACACGTTTCATCGATCCACCTTTCACGCTTTCGCTTGTATCCAAACATG 328
Qy 122 InIleAlaGlyAsnPheAlaProValAsnGluGlnProValArgArgAsnLeuProVal 142
Db 329 TCTCTCTGTAACTTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 388
Qy 142 al---GlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyAlaAsnProL 161
Db 389 TCCACGGCACCTTCTCCACTGTACCTTAACGGCGCTTACATCGTAAACGGTCCAAATCCAC 448
Qy 161 euHisGluProValThrGlyHisHisPheAspGlyAspGlyMetValHisAlaValL 181
Db 449 AGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 508
Qy 181 ysPheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheValG 201
Db 509 AAATCCCAACGGTAAAGCCACTCTCTGTAGCAGATACGCTCAAGACTTATATAATCAACG 568
Qy 201 InGluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGlyHis 221
Db 569 TCGAGAAACAAACCGGAGCTCCGGTTATGCTTAACGCTGTTTCCCGATTCACCGGTAA 628
Qy 221 hr---GlyIleAlaArgLeuMetPheTyrAlaArgAlaAlaAlaAlaAlaAlaAla 240
```

```
Db 629 CGGCGTCAGTAGCTCGTGAGGCTTTAACGGCAGCTAGGTTTAAACCGCAGCAGATATAATC 688
Qy 240 roAlaHisGlyThrGlyValAlaAenAlaGlyLeuValTyrPheAsnGlyArgLeuLeuA 260
Db 689 CGGTTAACGGCATTCGGTTAGCTATAACAAGCTAGCTTCTTCAGTAACCGCTCTCTTTG 748
Qy 260 laMetSerGluAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLeuYst 280
Db 749 CTTTAGGTGAATCTGATTACCTACGGCGTCGGAATTAACGAATCAGGAGATATTGAAA 808
Qy 280 hrValGlyArgPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysV 300
Db 809 CGATCGGACGGTACGATTTCGACGGAAATAGCGATGAGTATGACAGCTCATCTCAAAA 868
Qy 300 alAspProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValSerLysProTyrL 320
Db 869 CCGATCCCAATAACCGGAGAAATTTTCGCTTTCCGGTACGGTCCGGT---CCACCGTTT 925
Qy 320 euLysTyrPheArgPheSerProAspGlyThrLysSerProAspValGluIle---GlnL 339
Db 926 TAACATATTTCCGGTTTGAATCCGCGCGGAAAAAACAAGAGACGTTCCGATATTCTCGA 985
Qy 339 euAspGlnProThrMetHisAspPheAlaIleThrGluAsnPheValValProA 359
Db 986 TGACGCTCCGTCGTTTCTCCATGACTTCGGATCACGAACGTCACGCGATTTCGCAG 1045
Qy 359 spGlnGlnValPheLys-----LeuProGluMetIleArgGlyGlySerProV 376
Db 1046 AGATTTCAGCTGGCATGAGGATGAACATGTTGGATTTCGGTTCTCGAAGTGGTTCTCCGG 1105
Qy 376 alValTyrAspLysAsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspS 396
Db 1106 TTGGTACTGATAACCGGAAAACTCAAGCGTTGGAGTGAATTCCTAAGTACCGCGGAGATG 1165
Qy 396 erSerAsnIleLysTrpIleAspAlaProAspCysPheCysPheHisLeuTrpAsnAlat 416
Db 1166 ACTCGGAGATGAATGGTTCGAAGTTCCTGGATTCAATATCATTCACCGTATTATGCTT 1225
Qy 416 rpGluGluProGluThrAspGluValValValIleGlySerCysMetThrProProAsps 436
Db 1226 GGGATGAAGATGATGAAACACGCGTCTGTTGATTGCACCGAATATTATGCTGATTGAAC 1285
Qy 436 erIlePheAsnGluSerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAenL 456
Db 1286 ATACTTTAGAGAGGATGAT---CTGGTTCATGCTTTCTGGAGAGGTGAAGATCGATC 1342
Qy 456 euLysThrGlyGluSerThrArgArgProIleIleSerAsnGluAspGlnGlnValAsnL 476
Db 1343 TCGTCACCGGGATTGTGAGACGTCATCCGATCTCAGCGAGG-----AATC 1387
Qy 476 euGluAlaGlyMetValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeuA 496
Db 1388 TCGATTTCGCTGTGATTAATCCGGCGTTTCTCGGGAGATGTAGCAGGTACGTTACGCGG 1447
Qy 496 laLeuAlaGluProTrpProLysValSerGlyPheAlaLysValAspLeuThrThrGlyG 516
Db 1448 CGATTGGAGATCCGATGCGGAAGATCTCCGGTGTGCTGAAGCTTGATGTCTAAAGAG 1507
Qy 516 lu-----ValLysLysHisLeuTyrGlyAspAsnArgTyrGlyGlyGluP 531
Db 1508 ATCGGGATGATTGTACGTTGGCCCGTAGAATGTACGGTTTCAGGTGTTACGCGCGAGAAC 1567
Qy 531 roLeuPheLeuProGlyGluGlyGly-----GluGluaspGluGlyTyrIleL 547
Db 1568 CGTTTTCGTAGCTAGGATCTCGTGAATCCGGAGCGGAGAGATGATGTTATGTGG 1627
Qy 547 euCysPheValHisAspGluLysThrTrpLysSerGluLeuGlnIleValAsnAlaValS 567
Db 1628 TGACGTATGTTCCACGATGAAGTACTGGAGAAATCGAAGTTTCTGGTGAAGCAGCTAAAT 1687
Qy 567 er-----LeuGluValGluAlaThrValLysLeuProSerArgValProTyrGlyPheH 585
Db 1688 CCGCGGAGCTTGAATCGTCGCGCGGTGAGGTTCCCGCGAAGGTTCCGTACGAGTTCC 1747
```

```
Qy 585 isGlyThrPheIleGlyAlaAspAspLeuAlaLys 596
Db 1748 ATGGGTATTGTCAAGGAAAGTGACCTTAATAAG 1782
RESULT 9
US-09-758-269-7
; Sequence 7, Application US/09758269
; Patent No. US20020104120A1
; GENERAL INFORMATION:
; APPLICANT: IUCHI, SATOSHI
; APPLICANT: KOBAYASHI, MASATOMO
; APPLICANT: SHINOZAKI, KAZUO
; TITLE OF INVENTION: TRANSGENIC PLANTS CARRYING NEOXANTHIN
; TITLE OF INVENTION: CLEAVAGE ENZYME GENE
; FILE REFERENCE: 3914-3
; CURRENT APPLICATION NUMBER: US/09/758,269
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: JP 2001-003476
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: JP 2000-010056
; PRIOR FILING DATE: 2000-01-13
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 1617
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1614)
US-09-758-269-7
Alignment Scores:
Pred. No.: 1,64e-97 Length: 1617
Score: 938.00 Matches: 207
Percent Similarity: 55.81% Conservative: 105
Best Local Similarity: 37.03% Mismatches: 199
Query Match: 29.78% Indels: 48
DB: 13
US-09-758-269-6 (1-599) x US-09-758-269-7 (1-1617)
Qy 64 LysGlnSerSerAsnSerProAlaIleValLysProLys---AlaLysGluSerAsn 82
Db 10 AAATCTCAGTGTGAGCAGCAGCATCTCTCAGTCCATCTTAGACCTCCAAAGGTTTCTCC 69
Qy 83 ThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaLeuAspAlaAlaGluGly 102
Db 70 TCGAAGCTTCTCGATCTTCTCGAGAGACTTGTTCAGGCTCATG-----TAC 114
Qy 103 PheLeuValSerHisGluLysLeuHisProLeuProLysThrAlaAspProSerValGln 122
Db 115 -----CACGATGCTTCTCCTCCTCCAC-----TAC 141
Qy 123 IleAlaGlyAsnPheAlaProVal---AsnGluGlnProValArgAsnLeuProVal 141
Db 142 CTCTCAGCAACTTCGCTCCATCCGTCATGAAACTCTCCCGTCAGGATCTCCCGTC 201
Qy 142 ValGlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyValAsnProLeu 161
Db 202 CATGATTTCTCCGAATGCTTGAATGTTGAGGGTTGGTCCAAACCCCAAG 261
Qy 162 HisGluProValThrGlyHisPhePheAspGlyAspGlyMetValHisAlaValLys 181
Db 262 TTTGATGTCGTGCGTGGATATCACTGGTTTTCATGAGATGGGATGATTCATGGGGTACG 321
Qy 182 PheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheValGln 201
Db 322 ATCAAGATGGGAAGCTACTTATGTTTCTCGATATGTTAAGACATCAGCTCTTAAGCAG 381
Qy 202 GluArgGlnLeuGlyArgProValPheProLysAlaIleGlyGluLeuHisGlyHisThr 221
```



Db 382 GAAGAGTTCTTCGGAGCTGCCAAATTCATGAAG---ATTGGTGACCTTAAGGGTTTTTC 438  
Qy 222 GlyIleAlaArgLeuMetLeuPheTyrAlaArgAlaAlaAlaGlyIleValAspProAla 241  
Db 439 GGATTGCTAATGCTCAATGCTCAACAGCTGAGAACGAAGCTCAAAATATTGGACAACACT 498  
Qy 242 HisGlyThrGlyValAlaAsnAlaGlyIleValTyrPheAsnGlyArgLeuLeuAlaMet 261  
Db 499 TATGGAATATGGAACCTGCCAATACAGCCTCGTATATACACCTGGAACAACTTCTACATTA 558  
Qy 262 SerGluAspLeuProTyrGlnValGlnIleThrProAsnGlyAspLeuLysThrVal 281  
Db 559 CAGGAGGCAGATAAGCCGTACGTATCAAAAGTTTGGAGATGGAGACCTGCAAACTCTT 618  
Qy 282 GlyA-gPheAspPheAspGlyGlnLeuGluSerThrMetIleAlaHisProLysValAsp 301  
Db 619 GGTATAATAGATTATGACAAGAGATTGACCCACTCTCTCACTGCTCACCCAAAGTTGAC 678  
Qy 302 ProGluSerGlyGluLeuPheAlaLeuSerTyrAspValValSerLysProTyrLeuLys 321  
Db 679 CCGGTTACGGGTGAATGTTTACATTCGGCTATTGCG---CATACGCCACCTTATCTCACA 735  
Qy 322 TyrPheArgPheSerProAspGlyThrLysSerProAspValGluIleGlnLeuAspGln 341  
Db 736 TACAGATTATCTCGAAGATGCGATTATGATGACCCAGTCCCAATTACTATATACAG 795  
Qy 342 ProThrMetMetHisAspPheAlaIleThrGluAsnPheValValProAspGlnGln 361  
Db 796 CCTATCATGATCATGATTTTCTTACTTACTGAGACTTATGCAATCTTCATGATCTTCT 855  
Qy 362 ValValPheLysLeuProGluMetIleArgGlyGlySerProValVal---TyrAspLys 380  
Db 856 ATGCATCTCAGGCCCAAGAAATGTTGAAGAGAAGAAATGATATATCTCATTTGATCCC 915  
Qy 381 AsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSerSerAsnIleLys 400  
Db 916 ACAAAAAGGCTCGTTTGGTGTCTTCACCGCTATGCCAAGGATGAACCTATGATTAGA 975  
Qy 401 TrpIleAspAlaProAspCysPheCysPheHisLeuTrpAsnAlaTrpGluGluProGlu 420  
Db 976 TGGTTTGAGCTTCCCACTGCTTTATTTCCACACGCCCAATGCTTTGGAA-----GAA 1029  
Qy 421 ThrAspGluValValIleGlySerCysMetThrProAspPhe----- 435  
Db 1030 GAGGATGAATGCTCTCTCATCTTGTCTGTGAGAAATCCAGATCTTCACATGTCAGT 1089  
Qy 436 SerIlePheAsnGluSerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsn 455  
Db 1090 GGGAAAGTGAAAGAAAACCTCGAAATTTTGGCAACGAACGTGACGAATGAGATTCAAC 1149  
Qy 456 LeuLysThrGlyLysThrArgArgProIleLeuSerAsnGluAspGlnGlnValAsn 475  
Db 1150 ATGAAACGGGCTCAGCTTCTCAAAAAAACTATCCGATCTGCG----- 1194  
Qy 476 LeuGluAlaGlyMetValAsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeu 495  
Db 1195 GTTGATTTTCCCAAGTCAATGAGTGTACACCGGAAAGAAACAGACAGATACGTATATGA 1254  
Qy 496 AlaLeuAlaGluProTrpProLysValSerGlyPheAlaLysValAsp----- 511  
Db 1255 ACAATTCTGGACAGTATCGAAAGGTTTACCGGAATCATCAAGTTTGTATCTCGATCAGAA 1314  
Qy 512 -----LeuThrThrGlyGluValLysLysHisLeuTyr----- 522  
Db 1315 GCTGAGACAGGAAAGATGCTGGAAGTAGAGGTATATCAAGGAATATATACCTG 1374  
Qy 523 GlyAspAsnArgTyrGlyGluProLeuPheLeuProGlyGlyGlyGlyGluAsp 542  
Db 1375 GGAAGAAGGCAGATATGGTTCAGAGCTATCTATGTTCCGCGTGACAGCAGAGAAGAC 1434  
Qy 543 GluGlyTyrIleLeuCysPheValHisAspGluLysThrTrpLysSerGluLeuGlnIle 562  
Db 1435 GACGGTTACTTGATATCTTTGTTTCATGATGAAACACAGAGGAATCATCGTGACTGTG 1494

Qy 563 ValAsnAlaValSerLeuGluValGlu-----AlaThrValLysLeuProSerArgVal 580  
Db 1495 ATAGACGCAAAACAATATGCTCGCTGAACCGGTGGCAGTGGAGCTGCCGCACAGGTC 1554  
Qy 581 ProTyrGlyPhePheHisGlyThrPheIleGlyAlaAspAspLeuAlaLysGlnValVal 599  
Db 1555 CCATATGCTTCCCATGCTTGTGTTTATACAGAGGAACAACCTCCAGGAACAACACTCTT 1611  
RESULT 10  
US-09-758-269-17  
; Sequence 17, Application US/09758269  
; Patent No. US20020104120A1  
; GENERAL INFORMATION:  
; APPLICANT: IUCHI, SATOSHI  
; APPLICANT: KOBAYASHI, MASATOMO  
; APPLICANT: SHINOZAKI, KAZUO  
; TITLE OF INVENTION: TRANSGENIC PLANTS CARRYING NEOXANTHIN  
; TITLE OF INVENTION: CLEAVAGE ENZYME GENE  
; FILE REFERENCE: 3914-3  
; CURRENT APPLICATION NUMBER: US/09758,269  
; PRIOR FILING DATE: 2001-01-12  
; PRIOR APPLICATION NUMBER: JP 2001-003476  
; PRIOR FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: JP 2000-010056  
; PRIOR FILING DATE: 2000-01-13  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 17  
; LENGTH: 1617  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1614)  
US-09-758-269-17  
Alignment Scores:  
Pred. No.: 2,14e-97 Length: 1617  
Score: 937.00 Matches: 207  
Percent Similarity: 55.81% Conservative: 105  
Best Local Similarity: 37.03% Mismatches: 199  
Query Match: 29.75% Indels: 48  
DB: 10 Gaps: 13  
US-09-758-269-6 (1-599) x US-09-758-269-17 (1-1617)  
Qy 64 LysGlnSerSerAsnSerProAlaIleValLysProLys---AlaLysGluSerAsn 82  
Db 10 AAATCTAGTGTGCGAGCATCATCTCAGTCCATCTCAGTCCATCTCAGCCCTCCAGGGTTTCTCC 69  
Qy 83 ThrLysGlnMetAsnLeuPheGlnArgAlaAlaAlaAlaLeuAspAlaAlaGluGly 102  
Db 70 TCGAAGTCTTCGATCTTCGAGAGACTTCTCAGTCAAGCTCATG----- 114  
Qy 103 PheLeuValSerHisGluLysLeuHisProLysThrAlaAspProSerValGln 122  
Db 115 -----CAGGATGCTTCTCTCCCTCTCCAC-----TAC 141  
Qy 123 IleAlaGlyAsnPheAlaProVal---AsnGluGlnProValArgAsnLeuProVal 141  
Db 142 CTCTCAGCAACTTTCGCTCCCATCCGTCATCGTGAATACTCTCCCGTCAAGGATCTCCCGCTC 201  
Qy 142 ValGlyLysLeuProAspSerIleLysGlyValTyrValArgAsnGlyAlaAsnProLeu 161  
Db 202 CATGATTTCTTCGGAATGCTTGAATGGTGAATTTGTAGGGTTGGTCCAAACCCCAAG 261  
Qy 162 HisGluProValThrGlyHisPhePheAspGlyAspGlyMetValHisAlaValLys 181  
Db 262 TTTGATGCTGCTCGTGGATATCACTGTTGATGGAGATGGAGATGATTCATGGGTAGCC 321  
Qy 182 PheGluHisGlySerAlaSerTyrAlaCysArgPheThrGlnThrAsnArgPheValGln 201



## RESULT 12

US-08-976-063C-1/c  
 ; Sequence 1, Application US/08976063C  
 ; Publication No. US20020182697A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Alexander Steinbuechel; Horst Priefert; Jurgen Rabenhorst  
 ; TITLE OF INVENTION: SYNTHETIC ENZYMES FOR THE PRODUCTION OF  
 ; TITLE OF INVENTION: CONFERYL ALCOHOL, CONFERYLALDEHYDE, FERULIC ACID, VANILLIN A  
 ; TITLE OF INVENTION: ACID AND THEIR USE  
 ; NUMBER OF SEQUENCES: 42  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SPRUNG KRAMER SCHAEFER & BRISCOE  
 ; STREET: 660 White Plains Road  
 ; CITY: Tarrytown  
 ; STATE: New York  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10591-5144  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB storage  
 ; COMPUTER: HP VECTRA  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: Wordperfect 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/976,063C  
 ; FILING DATE: 21-NOV-1997  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 196 49 655.1 (Germany)  
 ; FILING DATE: 29-NOV-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Kurt G. Briscoe  
 ; REGISTRATION NUMBER: 33,141  
 ; REFERENCE/DOCKET NUMBER: Bayer 9998-CAO  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (914) 332-1700  
 ; TELEFAX: (914) 332-1844  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 32679 base pairs  
 ; TYPE: Nucleic Acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (Genomic)  
 ; HYPOTHETICAL: NO  
 ; ANTI-SENSE: NO  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Pseudomonas sp.  
 ; STRAIN: HR199  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 3146..3997  
 ; OTHER INFORMATION: /gene= "ORF1"  
 ; US-08-976-063C-1

Alignment Scores:  
 Pred. No.: 1,12e-35 Length: 32679  
 Score: 419.00 Matches: 160  
 Percent Similarity: 37.97% Conservative: 83  
 Best Local Similarity: 25.00% Mismatches: 229  
 Query Match: 13.30% Indels: 168  
 DB: 8 Gaps: 24

US-09-758-269-6 (1-599) x US-08-976-063C-1 (1-32679)

QY 19 HisThrGlnProLeuSerSerGlnSerSerAspLeuSerTyrCysSerSerLeu 38  
 |||||  
 Db 17011 CATACCGCGCCGGAATAGTTGTCATTCTTGAATTTAAACAATAGCGCGCTCGGC 16952  
 |||||  
 QY 39 ProMetAlaSerArgValThrArgLysLeuAsnValSerSerAlaLeuHisThrProPro 58  
 |||||  
 Db 16951 CCATTTTCA-----ATACCACCA 16934

QY 59 AlaLeuHisPheProLysGlnSerSerAsn-----SerProIleValVal 74  
 |||||  
 Db 16933 CCATTGCACCTCACAGACATCTTCACTTCAACCGCACACATGAATCCG-----GTC 16883  
 |||||  
 QY 75 LysProLysAlaLysGluSerAsnThrLys-----GlnMetAsnLeu 88  
 |||||  
 Db 16882 AGTCCTAAACATAAACTCCAAACCGCGCAAGCGCGCATTTAGCGCAATTCAGAACACAA 16823  
 |||||  
 QY 89 PheGlnArgAlaAlaAlaAlaAlaLeuAspAlaAlaGluGlyPheLeuValSerHisGlu 108  
 |||||  
 Db 16822 CGAGACCGTGCCATGGCGAGATTC----- 16799  
 |||||  
 QY 109 LysLeuHisProLeuProLysThrAlaAspProSerValGlnIleAlaGlyAsnPheAla 128  
 |||||  
 Db 16798 -----AACCGCAACGACCG-----CAATTAGTAGGAACACTTCTC 16763  
 |||||  
 QY 129 ProValAsnGluGlnProValArgAsnLeuProValValGlyLysLeuProAspSer 148  
 |||||  
 Db 16762 CCCACCGGTATAGAGCGAGACTTGTTCGATCTAGAGGTTGACGCGCAATTCCTCAATCA 16703  
 |||||  
 QY 149 IleLysGlyValTyrValArgAsnGlyAlaAsnProLeuHisGluProValThrGlyHis 168  
 |||||  
 Db 16702 ATAAATGGAACTTCTACCGTATATACCGCAGACCTCAAGTTACCCACAAAAATTCAC 16643  
 |||||  
 QY 169 HisPheAspGlyAspGlyMetValHisAlaValLysPheGluHisGlySerAlaSer 188  
 |||||  
 Db 16642 ACCTTCATAGATGGAGATGGAATGCCTCTGCCTTCCACTTCGAAGATGTCATGTCGAC 16583  
 |||||  
 QY 189 TyrAlaCysArgPheThrGlnThrAsnArgPheValGlnGlnArgGlnLeuGlyArgPro 208  
 |||||  
 Db 16582 TTCATCAGTCGCTGGGTTAAAAACCGTCGATTCACGCGCGAAGCACTAGCGGAAAAATCG 16523  
 |||||  
 QY 209 ValPheProLysAlaIleGlyGluLeuHisGlyHisThrGlyIleAlaArgLeuMetLeu 228  
 |||||  
 Db 16522 CTATTGGCATGTACAGAAACCCCTATACCGAGACACAGTGTAAAGGACTA----- 16469  
 |||||  
 QY 229 PheTyrAlaArgAlaAlaAlaGlyIleValAspProAlaHisGlyThrGlyValAlaAsn 248  
 |||||  
 Db 16468 -----GACCGCACCGTGTCCAAT 16451  
 |||||  
 QY 249 AlaGlyLeuValTyrPheAsnGlyArgLeuLeuAlaMetSerGluAspAspLeuProTyr 268  
 |||||  
 Db 16450 ACAAGCATCATTAGCCATCACCGCAAGTGTCTGGCGGTGAAGAAAGACGCGCTACGCTAC 16391  
 |||||  
 QY 269 GlnValGlnIleThrProAsnGlyAspLeuLysThrValGlyArgPheAspPheAspGly 288  
 |||||  
 Db 16390 -----GAACCTGGATCCTCGT---ACACTTGAACCTCGCGGACACTTCGACTACGCGC 16340  
 |||||  
 QY 289 GlnLeuGluSer---ThrMetIleAlaHisProLysValAspProGluSerGlyGluLeu 307  
 |||||  
 Db 16339 CAAGTTACCGCAACCAACCCACACCGCCATCCAAAATATGACCCCGAAACGGGTACTTG 16280  
 |||||  
 QY 308 PheAlaLeuSerTyrAspValValSerLysProTyrLeuLysTyrPheArgPheSerPro 327  
 |||||  
 Db 16279 TTG-----TTCTTCGTTTCGCGAGCT 16259  
 |||||  
 QY 328 AspGlyThrLysSerProAspVal----- 335  
 |||||  
 Db 16258 AAGGCGGAAGCAACTCCAGACATGCGCTATTATCATTTGTGCAACACGCGAAGGTGACA 16199  
 |||||  
 QY 336 ---GluIleGlnLeuAspGlnPro-----ThrMetMetHisAspPheAlaIleThrGlu 352  
 |||||  
 Db 16198 CATGAACCTTGGTTTGAGCAGCGCCTTGGCGCATTCATGCACGACTTGCCTATACCCGA 16139  
 |||||  
 QY 353 AsnPheValValProAspGlnGlnValValPheLysLeuProGluMetIleArgGly 372  
 |||||  
 Db 16138 AATTGGTCCATTTTCCCAATTATGCGCGCACCAACAGCGCTG---TCCCGCTCAAGGCG 16082  
 |||||  
 QY 373 GlySerProVal---ValTyrAspLysAsnLysValAlaArgPheGlyIleLeuAspLys 391  
 |||||  
 Db 16081 AAACGCAATTTATATGTGGAGCGCGAACTGGGCAGCTACATTTGGCGGTACTC----- 16028



```
Db 1080 CTTGAACCTCGCGACACTTCGACTAGCAGCGCCAGTTACAGCCAAACCCACACCGCC 1021
QY 297 HisProLysValAspProGluSerGlyGluPheAlaLeuSerTyrAspValValSer 316
Db 1020 CATCAAAATATGACCGCGAAACGGGTGACTTGTG----- 985
QY 317 LysProTyrLeuLysTyrPheArgPheSerProAspGlyThrLysSerProAspVal--- 335
Db 984 -----TTCTTCGGTTCGCGAGCTAAGGCGGAAGCAACTCCAGACATGGCC 940
QY 336 -----GluileGlnLeuAspGlnPro--- 342
Db 939 TATTACATTGTCGACAGCAGCGCAGGTGACATGAACACTGGTTGACGAGCCCTAT 880
QY 343 ---ThrMetMetHisAspPheAlaIleThrGluAsnPheValValValProAspGlnGln 361
Db 879 GCGCATTCATGACGACTTGTGCATTACCGAAATTTGTCATTTCCTCAATTATGCGG 820
QY 362 ValValPheLysLeuProGluMetIleArgGlyGlySerProVal---ValTyrAspLys 380
Db 819 GCCACCAACAGCGCTG---TCCCGCCTCAAGCGGAAACAGCCCAATTTATATGTGGAGCG 763
QY 381 AsnLysValAlaArgPheGlyIleLeuAspLysTyrAlaGluAspSerSerAsnIleLys 400
Db 762 GAATCGGCGAGTACATTGCGGTACTC-----GCCCGCGCGAGGCGAGTCTGATTGCG 709
QY 401 TrpIleAspAlaProAspCysPheCysPheHisLeuTrpAsnAlaTrpGluProGlu 420
Db 708 TGGCTCAAGGCCACCGCGCTCTGGGTATTTCATGTTGTAATGCTTGGGAAGTCGGAACC 649
QY 421 ThrAspGluValValIleGlySerCysMetThrPro-----ProAspSer--- 436
Db 648 AAGATTATATGACCTTATGAAAGTGAATCTCGCGTTCCTCCCTTCCCAACTCACAA 589
QY 437 -----ile 437
Db 588 AACCAACCTTCGCCCTCAGAAAGCCGTACACGCGCTGACTCGTTGGGAAATTCACCTC 529
QY 438 PheAsnGluSerAspGluAsnLeuLysSerValLeuSerGluIleArgLeuAsnLeuLys 457
Db 528 GATAGCAGCAGCGAGATCAAGCGAACCCGCTACACGATTCTTTCGCGAAATGCCA 469
QY 458 ThrGlyGluSerThr----- 462
Db 468 ATCATGGATTCTTCGTCCTCCCTGCAATGCAACCGCTATGCTGTTATGGGGTGGACCAT 409
QY 463 ---ArgArgProIleIleSerAsnGluAspGlnGlnValAsnLeuGluAlaGlyMetVal 481
Db 408 CCACGCAAAACCACTTCGCGCATCAGCAGCGCGAGAAGATA----- 370
QY 482 AsnArgAsnMetLeuGlyArgLysThrLysPheAlaTyrLeuAlaLeuAlaGluProTrp 501
Db 369 -----TTCCGCTACAACCTCACTCGGC---ATCTGG 343
QY 502 ProLysValSerGlyPheAlaLysValAspLeuThrThrGlyGluValLysLysHisLeu 521
Db 342 -----GACAACACCACCGAGGTGACTACCACTCTGGTAC 310
QY 522 TyrGlyAspAsnArgTyrGlyGlyGluProLeuPheLeuPro---GlyGluGlyGlyGlu 540
Db 309 TCCGCGCAAGCCTCGCGCGCCCGCAGAGCGCGCTTCGTCCTCGATGAAATCGCAGCATCTG 250
QY 541 GluAspGluGlyTyrIleLeuCysPheValHisAspGluLysThrTrpLysSerGluLeu 560
Db 249 GAAGGTGAGGGTACTTGTGACCGGTGTGCTCGCTCGATGAAATCGCAGCATCTG 190
QY 561 GlnIleValAsnAlaValSerLeuGlu-----ValGluAlaThrValLysLeuProSer 578
Db 189 GTAATTCGCACTCAAGACATCCAGTCTGGTCCGTCGCAACCATCAAGTCCGCTC 130
QY 579 ArgValProTyrGlyPheHisGlyThrPheIle 589
Db 129 CGGCTAAGGCGCGCTCTCCATGGCTGCTGGTA 97
```

## RESULT 14

```
US-09-878-574-2872
; Sequence 2872, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 2872
; LENGTH: 320
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(320)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: LIB3028-019-Q1-B1-B11
US-09-878-574-2872
```

## Alignment Scores:

Pred. No.:	4,598-33	Length:	320
Score:	369.00	Matches:	68
Percent Similarity:	80.73%	Conservative:	20
Best Local Similarity:	62.39%	Mismatches:	15
Query Match:	11.71%	Indels:	6
DB:	10	Gaps:	2

US-09-758-269-6 (1-599) x US-09-878-574-2872 (1-320)

```
QY 447 SerValLeuSerGluIleArgLeuAsnLeuLysThrGlyGluSerThrArgArgProIle 466
Db 2 AGCGTTTAAACAGAGTAGGCTGAAACATGAGAACGGGAGGAGGGTGTG 61
QY 467 IleSerAsnGluAspGlnGlnValAsnLeuGluAlaGlyMetValAsnArgAsnMetLeu 486
Db 62 GTG-----GAGGAAATGAACCTGGAGGCAGGATGTTGAAACAGGAAAGTTG 109
QY 487 GlyArgLysThrLysPheAlaTyrIleAlaLeuAlaGluProTrpProLysValSerGly 506
Db 110 GGGAGAAAACACACGGTTCCGATATTTGTGCATAGCGGAACCGTGGCCGAAAGTGTGGGG 169
QY 507 PheAlaLysValAspLeuThrThrGlyGluValLysLysHisLeuTyrGlyAspAsnArg 526
Db 170 GTGGCGAAGGTGGACCTGGAGAGTGGGGAAGTGAAGACAGGCAGATACGAGAGAGAAG 229
QY 527 TyrGlyGlyGluProLeuPheLeuPro-----GlyGluGlyGlyGluAspGluGly 544
Db 230 TTCGGTGAGAGCCATTCTTCTTCCACACACCTGCGAGGGAATGGGAATGAGCATGAAGG 289
QY 545 TyrIleLeuCysPheValHisAspGlu 553
Db 290 TACGTGATGGCTTTTGTGCATGACGAG 316
```

## RESULT 15

```
US-09-770-696-172/c
; Sequence 172, Application US/09770696
; Patent No. US2001004940A1
; GENERAL INFORMATION:
; APPLICANT: Goriach, Jörn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
```

